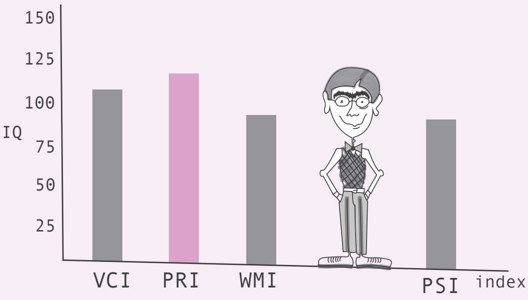


Table 3. Goodness of Fit for 2,3 and 4 Factor Models

Factors	X^2	df	p
2	191.33	89	.000
3	129.09	75	.000
4	68.38	62	.270

Note. df = Degrees of Freedom

The relationship between intelligence and executive function

Understanding theory in clinical practice

Loes van Aken

Table 1. Demographics of the Sample Population

	N	% Male	Age (years)		Total IQ	
			M	SD	M	SD
Total	188	51.6	39.5	15.5	93.4	17.9
Patients	138	56.5	41.6	15.1	88.0	15.9
Healthy participants	50	38.0	33.5	15.2	111.9	10.8

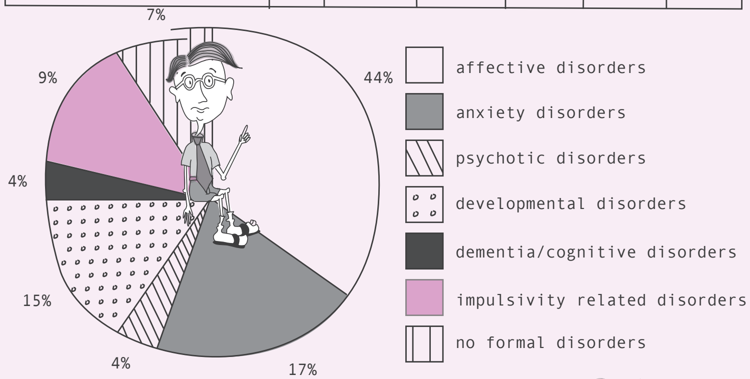
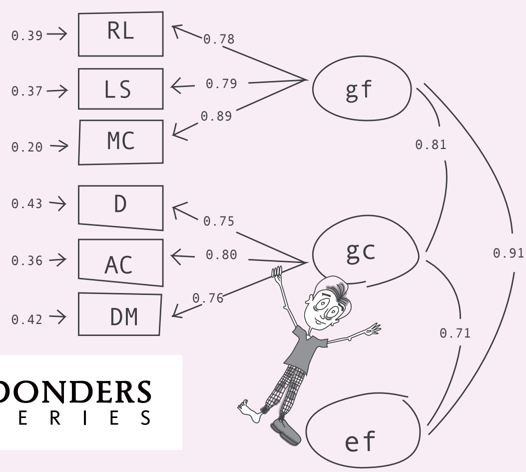
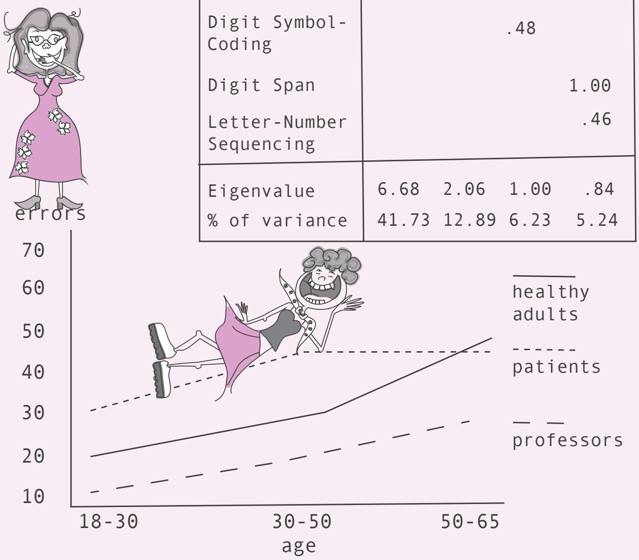


Table 4. Rotated Four Factor Solution

	Factors			
Test	I	II	III	IV
Vocabulary	1.02			
Comprehension	.92			
Information	.80			
Similarities	.65			
Arithmetic	.51			
Matrix Reasoning		.84		
Picture Completion		.72		
Picture Arrangement		.67		
Block Design		.66		
BADS		.62		
WCST		-.39		
Stroop		-.35		
Symbol Search			.99	
Digit Symbol-Coding			.48	
Digit Span				1.00
Letter-Number Sequencing				.46
Eigenvalue	6.68	2.06	1.00	.84
% of variance	41.73	12.89	6.23	5.24

Figure 1. Final model





DONDERS SERIES

ISBN
978-94-6284-094-2

Cover
Lieke Linssen

Design/lay-out
Promotie In Zicht, Arnhem

Print
Ipskamp Printing, Enschede

The research of this thesis was supported by the Vincent van Gogh Institute for Psychiatry, Venray, the Donders Institute for Brain, Cognition, and Behaviour, Centre for Cognition, Radboud University, Nijmegen, and the Behavioural Science Institute, Radboud University, Nijmegen, the Netherlands.

The relationship between intelligence and executive function

Understanding theory in clinical practice

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens het besluit van het college van decanen
in het openbaar te verdedigen op vrijdag 12 mei 2017
om 10.30 uur precies

door

Loes van Aken
geboren op 9 augustus 1987
te Venlo

Promotoren
Prof. dr. J.I.M. Egger
Prof. dr. R.P.C. Kessels

Copromotor
Dr. P.T. van der Heijden

Manuscriptcommissie
Prof. dr. C.L.M. Witteman (voorzitter)
Prof. dr. W.C.M. Resing (Universiteit Leiden)
Prof. dr. J.M. Bouma (Rijksuniversiteit Groningen)

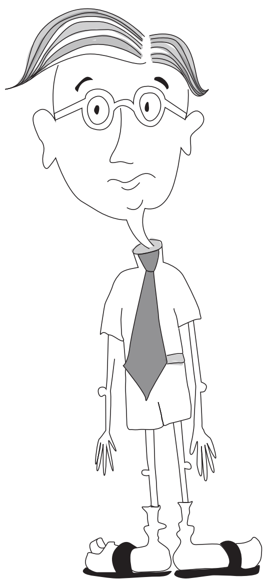
Een statisticus waadde vol vertrouwen door een rivier
die gemiddeld één meter diep was. Hij verdronk.

Godfried Bomans (1913-1971)

Contents

Chapter 1

Introduction



In clinical practice, intelligence testing is a common part of (neuro) psychological assessment. The intelligence quotient (IQ score) tends to correlate positively with performance on any other neuropsychological test. This phenomenon is also known as the *positive manifold*. Especially the executive functions (EF) show a large amount of overlap with the construct of intelligence (Chuderski, 2013; Diamond, 2013; Duggan & Garcia-Barrera, 2015; Duncan, Schramm, Thompson, & Dumontheil, 2012; Godoy, Dias & Sewabra, 2014; Redick, Unsworth, Kelly & Engle, 2012; Salthouse, Atkinson & Berish, 2003; Salthouse & Pink, 2008). Looking at the resemblance of both IQ tests and executive tests, this should not be much of a surprise. Despite their similarities, however, they have very different historical and theoretical backgrounds.

The main conceptualization of intelligence arises from a psychometric perspective. Having origins in the psychology of individual differences and academic psychology, intelligence testing mainly takes place within a normal and healthy population, although intelligence can also be part of clinical assessment. Using factor analyses, the aim in psychometrics is to identify different cognitive components contributing to intelligent behaviour, comprehensively described in the Cattell-Horn-Carroll theory of cognitive abilities (CHC theory; McGrew, 2009; Schneider & McGrew, 2012).

EF originates from neuropsychological theory. It concerns 'higher cognitive functions', responsible for goal-oriented and efficient behaviour in new and complex situations. The neuropsychological approach focuses on cognitive dysfunction in brain-injured or psychiatric patients. Furthermore, the study of EF entails unraveling the underlying neurocognitive processes. Thus, the concept of EF is embedded as a key element both in neuropsychological information processing models as well as in patient care.

Despite their different origins, there is ample evidence in support of the notion that intelligence and EF share great similarities at a conceptual level. However, the degree of overlap between the two constructs remains not fully clear, and clarification of theory and operationalization of the constructs is needed.

The main aim of the current thesis is to reach better understanding of (1) the relation between intelligence and EF at a conceptual level following current theories on both constructs, and (2) how intelligence and EF are measured and utilized in neuropsychological assessment. In the following introduction, a brief description will be given of the background of both theory and operationalization of intelligence and EF. Subsequently, the structure of this dissertation will be specified in the thesis outline.

Theories of intelligence

The concept of intelligence goes back to Darwin's ideas on inheritance, where individual differences were considered of high importance in the process of natural selection. His cousin Francis Galton elaborated on this idea, seeking to understand how differences between (groups of) individuals evolved. He observed a general tendency of 'well-doing' or success in daily life (later identified as the positive manifold), which Spearman subsequently defined as *g* or 'general intelligence' (Spearman, 1927). Even though the concept of intelligence is over a hundred years old, the definition and terminology of ('general') intelligence is still widely discussed. Although no consensus about the definition of intelligence is reached (yet), the terminology is used on an everyday basis. The absence of a consensus definition hampers the study of intelligence in current social sciences, which is further complicated as most scholars reject the idea that intelligence is a single entity describing an individual's level of functioning. At the same time, after a century of research, generic factors pop up in every dataset, reflecting the overall level of performance, and intelligence tests are still widely used in clinical assessments.

The concept of intelligence is studied using many different approaches, resulting in different levels of interpretation (see Gardner, 2012; Kaufman, Kaufman & Plucker, 2013, for a brief overview). For instance, intelligence is described from the perspective of biological and cognitive models, concerning cellular functioning, neural efficiency or networks, and information processing theories on cognitive functioning. Psychometric (trait) models are concerned with the evaluation of individual differences, looking for common cognitive components in large groups of data. There is also a behavioural perspective on intelligence, although behavioural analysts basically reject the idea of an intelligence construct being the cause of intelligent behaviour, particularly problem solving.

Research in all these different fields resulted in an array of intelligence theories. It goes beyond the scope of this thesis to discuss all these different approaches at length, but it is important to note that many of these theories concerning biology and brain networks have not led to adequate operationalization yet, limiting their clinical usefulness. Psychometric theories on intelligence on the other hand, have resulted in the successful translation into applicable measurements, which are in turn most frequently used in applied research and clinical practice. In the following section I will therefore take a closer look at these psychometric models of intelligence.

Psychometric or trait models

The general factor *g* was identified using factor-analytic techniques on large sets of data. It did not take long for researchers, however, to conclude that more than one single factor was responsible for defining all human capacities. First, a distinction between fluid intelligence (Gf) and crystallized intelligence (Gc) was proposed by Horn and Cattell (1966). This dichotomy distinguishes between the ability to solve complex and new problems using reasoning, independent of prior knowledge (Gf),

and a knowledge based ability that concerns the depth and breadth of acquired knowledge or experience through schooling and acculturation (Gc). Later, Carroll (1993) reanalyzed data of over 460 datasets and created his hierarchical three-stratum model, describing the general factor (stratum III) on top of the hierarchy, eight broad abilities including Gf and Gc (stratum II), and multiple narrow abilities (stratum I) at the bottom or the hierarchy. Extension of the Gf-Gc theory (Horn & Cattell, 1966) combined with Carroll's three-stratum model (Carroll, 1993) eventually led to what is now referred to as the Cattell-Horn-Carroll theory of cognitive abilities (CHC theory; McGrew, 2009; Schneider & McGrew, 2012).

The CHC model can be used as a taxonomy of (potentially) all existing cognitive functions. The three strata used by Carroll (1993) are at the core of the CHC model, and currently around 16 broad abilities (comparable with Stratum II) have been identified (Schneider & McGrew, 2012). Furthermore, the CHC model distinguished six domain-independent general capacities: fluid reasoning (Gf), short-term memory (Gsm), long-term storage and retrieval (Glr), processing speed (Gs), reaction and decision speed (Gt), and psychomotor speed (Gps). There are four abilities that address acquired knowledge; comprehension-knowledge (Gc), domain-specific knowledge (Gkn), reading and writing (Grw), and quantitative knowledge (Gq). Six abilities concern domain specific sensory and motor functions (visual processing; Gv, auditory processing; Ga, and olfactory, tactile, kinesthetic and psychomotor abilities; Go, Gh, Gk, and Gp).

Theories of executive functioning

The neuropsychological approach of the study of human behaviour focuses on the differentiation between normal and pathological cognitive functioning. Studying cognition, neuropsychologists strive for a detailed description of (1) personal and emotional variables, (2) cognitive functioning, and (3) executive functioning (Lezak, Howieson, Bigler & Tranel, 2012). These variables can be examined in various settings, for instance rehabilitation, school psychology, psychiatry, neurology, and neurosurgery. Cognitive functions refer to specific processes, including receptive functions, memory and learning, expressive functions, thinking, and mental activity or functions concerning the efficiency of mental processes. 'Higher order' cognitive functions, responsible for processes such as planning, monitoring, controlling, manipulation, and goal-directed behavior, are defined by the term executive functioning. In general, nine executive processes can be identified: attention, emotion regulation, flexibility, inhibitory control, initiation, organization, planning, self-monitoring, and working memory (Goldstein, Naglieri, Princiotta & Utero, 2014; Naglieri & Goldstein, 2013).

Information processing models

An important issue in understanding EF is the unity versus diversity debate, in which working mechanisms of EF are described as either distinct functions or constituting a controlling entity functioning as a unitary system (Miyake, Friedman, Emerson, Witzki, Howerter & Wager, 2000). The general consensus nowadays is that both unity and diversity in EF can co-exist. Preference for one or the other is mostly dependent on context; in cognitive neuroscience EF is mostly handled as a unitary construct, whereas in clinical neuropsychology the diversity of EF is accentuated (Duggan & Garcia-Barrera, 2015; Eling, 2016). EF as a unitary system is described in the widely used working memory model by Baddeley (2003, 2007). He identified EF as a central executive system, which comes into action when a situation requires more than automatic responses. The central executive is responsible for selecting and organizing behavior, shifting attention between actions and inhibition of actions. While doing so, the central executive uses temporally stored information that is held active within domain-specific slave-systems for visual, auditory or phonological information. The central executive is somewhat similar to the Supervisory Attentional System (SAS) described by Norman and Shallice (1986). The SAS functions as a selection-tool, deciding which actions have priority in complex situations which require goal-oriented behavior.

Emphasizing on the diversity of EF, Miyake et al. (2000) identified three core building blocks or sub processes of EF using factor analysis; shifting, updating, and inhibition of prepotent responses. These components are incorporated in an overview by Diamond (2013). She describes working memory, inhibitory control and cognitive flexibility as executive components, as well as reasoning and problem-solving, which in turn are synonymous to fluid intelligence. Naglieri & Goldstein (2013) consider EF to be a single phenomenon describing task efficiency and problem solving across nine areas: attention, emotion regulation, flexibility, inhibitory control, initiation, organization, planning, self-monitoring, and working memory (Goldstein et al., 2014; Naglieri & Goldstein, 2013). As Müller, Langner, Cieslik, Rottschy & Eickhoff (2015) state, executive functioning is an interplay of different executive processes. Dependent on specific task demands, the contribution of each process varies, but the resources are alike in every EF task (Müller et al., 2015). In this way, EF processes can be seen as tools which interact as a transcending system for the execution of complex behaviour.

In cognitive neuroscience, including neuropsychology, terminology like working memory, cognitive control, executive control, processing speed and perception is used to describe complex cognitive processes, or intelligent behavior. Neuropsychologists are concerned with clarifying cognitive dysfunctioning, studying (deficits in) behaviour using validated neuropsychological tests. Although neuropsychology does not revolve

around a single concept like *g* (or a full-scale IQ score for that matter) to explain human behaviour, intelligence testing is daily practice in neuropsychological assessment.

Current directions

Although in the past decades no other model of cognitive abilities has been studied more extensively than CHC (Alfonso, Flanagan & Radwan, 2005; Schneider & McGrew, 2012), some shortcomings become evident. Some research is being done examining the role of EF in relation to CHC, but EF does not have an explicit place within this model yet. This is remarkable, given the strong conceptual overlap and the high correlations seen between EF and the psychometric concept of Gf and Gsm (Duncan, 2013; Duncan, Schramm, Thompson & Dumontheil, 2012; Friedman, Miyake, Corley, Yung, DeFries & Hewitt, 2006; Roca et al., 2010; Salthouse, Atkinson & Berish, 2003; Salthouse & Davis, 2006; Salthouse & Pink, 2008). Factor analytic studies looked into this matter, relating tasks of EF mainly with CHC abilities Gf, Gv, and Gs (Floyd, Bergeron, Hamilton & Parra, 2010; Hoelzle, 2008, Jewsbury, Bowden & Duff, 2016; Roberds, 2015; Salthouse, 2005). The absence of EF processes creates a gap in CHC theory, and integration with neuropsychology is needed to transform CHC from a mere taxonomy of abilities into an information processing theory on human cognition.

Looking beyond the scope of CHC, some attempts are made to integrate psychometric intelligence theory with information processing theories. For instance, the multiple-demand system (MD system) by Duncan (2013) offers an alternative view on the positive manifold in terms of cognitive processing, based on neuroimaging findings. Regardless whether a task or situation requires planning skills, memory, language, or other cognitive abilities, similar brain regions of the MD system become active. Core job of the MD system is 'to control complex behaviour in a series of attentional episodes' (p. 37). Instead of having specific brain regions responsible for specific cognitive demands, activity in the MD system depends on the complexity or novelty of any cognitive task. Duncan (2013) states that this MD activity, or the efficiency with which novel/complex tasks of any kind are solved, is the core aspect of the psychometric concept of Gf. This could be an explanation of the positive manifold, creating positive correlations between independent demands due to the same underlying MD system.

In sum, CHC theory originates from a psychometric point-of-view, describing a taxonomy of separate cognitive abilities in (mainly) healthy individuals. The construct of EF has its roots in neuropsychological theory and is primarily described and examined from an information processing perspective on brain injury or psychiatric disease. Whereas the psychometric approach aims to establish internal and structural validity of measures to fully and adequately map the whole spectrum of cognitive abilities, the neuropsychological approach has its focus on predictive and external validity of measures, aiming to distinguish between normal and pathological behaviour. The next section will focus on the operationalization of these constructs, and how they are measured in clinical practice.

From theory to practice: Measurement of intelligence and EF

Although much has been written about intelligence and EF, clinical practice always lags behind theoretical developments. This holds especially for intelligence testing. CHC components can be identified in current intelligence tests, but tests based on CHC theory as a theoretical framework like the Woodcock Johnson Battery – Third Edition (Woodcock, McGrew & Mather, 2001) are not widely used and not even available in the Netherlands. Even though CHC plays a significant role in the development of the latest versions of intelligence tests like the Wechsler Adult Intelligence Test – Fourth Edition (Wechsler, 2008), the basis of these instruments is mainly empirical and pragmatic, instead of relying on theoretical frameworks describing processes of cognition.

In neuropsychology, particularly in the domain of EF, ill-defined concepts and overlap between constructs create difficulties in the operationalization of this cognitive construct (Cox et al., 2014). Any scientist who is concerned with EF research is troubled by the task impurity problem (Miyake et al., 2000). Most available tasks do not only measure the cognitive process the researcher is interested in, but also rely on other abilities, which may be nonspecific, such as motor function, non-executive attention or processing speed, or overlapping with other executive processing (e.g., successful task switching also requires inhibition of the ongoing task). Disentangling these processes is often complicated, especially in a clinical context. Consequently, if a person fails on an EF test, it is unknown which specific process is responsible for this failure. Failing to successfully perform a planning task, for example the Tower of London, could be a result of inadequate planning, lack of overview, poor inhibitory skills, or losing the attentional focus, to name a few. Traditional tests of EF can, in terms of the MD system, be fully explained by Gf (Duncan, 2013). This could (but does not necessarily) mean that EF is isomorphic with Gf. Another explanation is that those conventional EF tests do not measure specific functions (outside the MD system) enough (Duncan, 2013).

Integration of CHC theory and EF: Objectives

In sum, theories on EF and intelligence do not easily translate into adequate measurements of both constructs. Although they stem from two separate fields of interest, neuropsychology and psychometrics, tests of EF and intelligence are used interchangeably without exact knowledge of how these constructs relate to one

another. It is not yet clear whether CHC theory provides a good framework for cognitive functioning in disabled populations (Schneider & McGrew, 2012), although first results are promising (Jewsbury et al., 2016). Taking a first step towards an integrated approach, this thesis primarily focuses on how the constructs of EF and intelligence are exactly represented in contemporary neuropsychological assessment and how they compare and relate to each other. Furthermore, it questions whether CHC as guiding model in current intelligence research is compatible with current neuropsychological practice.

Thesis outline

Gaining insight into daily used instruments of intelligence and EF by adopting a pragmatic approach is the aim of **Chapter 2**. Using exploratory factor analysis, we examine whether EF is incorporated in the dominant intelligence test of the past decades, the Wechsler Adult Intelligence Test – third edition (WAIS-III; Wechsler, 1997a, 1997b). WAIS-III subtest performance was compared to performance on widely-used executive tests, that is, the Wisconsin Card Sorting Test, the Behavioural Assessment of the Dysexecutive Syndrome, and the Stroop Color Word Test.

The representation of the CHC theory in the latest Dutch-language version of the Wechsler scales, the WAIS-IV-NL (Wechsler, 2012a, 2012b), is examined and discussed in **Chapter 3**. Using confirmatory factor analysis, a five-factor structure according to CHC next to the four indices of the manual is examined in a psychiatric population.

Central to **Chapter 4** is the relation between Gf and EF. This specific relationship will be examined within a model based on the Gf-Gc dichotomy. A latent factor model is developed to test interrelations between Gf, Gc, and EF. The Kaufman Adolescent and Adult Intelligence Test (KAIT; Dekker, Mulder & Dekker, 2005; Kaufman & Kaufman, 1993) is used as an established measure of both Gf and Gc. EF is operationalized through various subtests of the Cambridge Neuropsychological Test Automated Battery (CANTAB).

Chapter 5 focuses on the gap between neuropsychological assessment and CHC theory, a problem which has emerged but is not solved through factor-analytic studies. The main question is whether the two can be integrated, examined in a large group of psychiatric patients. Six executive tasks (the Wisconsin Card Sorting Test, Stroop, Trailmaking Test, Rey's Complex Figure Test, Verbal Fluency, and the Tower of London), and multiple intelligence tests (the WAIS-III, WAIS-IV, and KAIT) are included in the study.

Finally, in **Chapter 6**, the results of the individual studies are integrated and discussed in terms of theoretical and clinical considerations. The chapter ends with an overall conclusion and suggestions for future research directions.

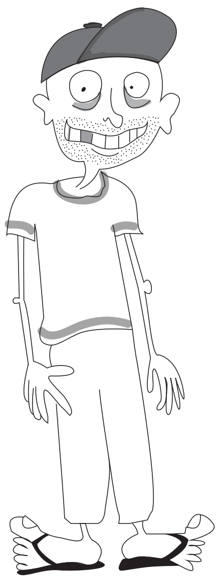
Chapter 2

Exploring the Incorporation of Executive Functions in Intelligence Testing: Factor Analysis of the WAIS-III and Traditional Tasks of Executive Functioning

Van Aken, L., Kessels, R.P.C., Wingbermühle, E., Wiltink, M., Van der Heijden, P.T., & Egger, J.I.M. (2014). Exploring the Incorporation of Executive Functions in Intelligence Testing: Factor Analysis of the WAIS-III and Traditional Tasks of Executive Functioning. *International Journal of Applied Psychology, 4,* 73-80. doi: 10.5923/j.ijap.20140402.05

Abstract

The aim of this study was to examine the relationship between subtests of the Wechsler Adult Intelligence Scale – third edition (WAIS-III) and executive functions. The Behavioural Assessment of the Dysexecutive Syndrome, Wisconsin Card Sorting Test, and Stroop Color-Word Test were administered to a heterogeneous group of 234 psychiatric patients and 24 healthy volunteers. Maximum likelihood procedures with promax rotation were applied to two, three and four factor solutions. The four factor model fit the data best, confirming the four factor indices of the WAIS-III. All three executive tasks had their highest loading on the factor corresponding to the perceptual organization index (POI) of the WAIS-III. Results confirm the overload of crystallized intelligence in the subtests and EF involvement in the POI of the WAIS-III. Results are discussed as to the need for an integrated, multifaceted view on cognitive disorders and intellectual (dis)abilities.



Introduction

The Wechsler Adult Intelligence Scale—Third Edition (WAIS–III; Wechsler, 1997a; Wechsler, 1997b) is a frequently used measure of intelligence. The development of subtests of the original WAIS occurred based on Wechsler's clinical experience. Empirical research led to modifications of the test, and factor analytical research on the subtests revealed a four factor structure. According to the recurrent findings, the WAIS-III structurally identifies four indices, i.e., the Verbal Comprehension Index (VCI), the Perceptual Organization Index (POI), the Working Memory Index (WMI), and the Processing Speed Index (PSI). These index scores provide better profile interpretation as compared to the verbal-performance (VIQ-PIQ) dichotomy (Kaufman & Lichtenberger, 1999). Moreover, they show a much better fit than the VIQ-PIQ factor solution in diverse clinical and nonclinical samples (Arnau & Thompson, 2000; Ryan & Paolo, 2001; Taub, 2001; Van der Heijden & Donders, 2003a; Van der Heijden, van den Bos, Mol & Kessels, 2013). Unfortunately, the WAIS-III held on to the dichotomy next to the four indexes, resulting in perseverance in the use of the VIQ and PIQ among clinicians. Therefore, the latest revision of the Wechsler Intelligence Scale, the WAIS-IV (Wechsler, 2008) does not provide VIQ and PIQ scores any longer.

Although structures were re-evaluated in the development of the WAIS-IV (mainly by eliminating VIQ and PIQ and replacing them with four index scores and a full scale IQ score), the WAIS-III is still widely used in clinical practice. Therefore, its structure and applicability to neuropsychological assessment should be reconsidered. The usefulness of the index scores, especially PSI and WMI, in neuropsychological evaluation is reasonably well established [Hawkins (1998); Martin, Donders and Thompson (2000); Fisher, Ledbetter, Cohen, Marmor & Tulsky, 2000; Taylor & Heaton (2001); Van der Heijden and Donders (2003b)], but the overall structure of the WAIS-III lacks theoretical ground. Therefore, the WAIS-III research findings have been subject to discussion within a framework of existing neuropsychological and factor analytical theories of intelligence (Ardilla, 1999; Duncan, 2010; Duncan, Burgess & Emslie, 1995; McGrew, 2009; Van der Heijden & Donders, 2003b).

One of the most influential theories is the Cattel-Horn-Carroll (CHC) theory of cognitive abilities. The CHC theory arose from the distinction between fluid (Gf) and crystallized (Gc) intelligence made by Horn and Cattell (1966) and Carroll's (1993) three striatum theory of cognitive abilities. The CHC theory consists of both a general component of intelligence (*g*; stratum III), broad abilities (stratum II, e.g. fluid reasoning, crystallized knowledge, visual and auditory processing, short-term memory, long- term storage retrieval, processing speed, decision and reaction speed, reading and writing and quantitative knowledge) and narrow abilities (stratum I), providing a complete and comprehensive taxonomy of human intelligence. See Kaufman, Kaufman & Plucker (2013), and McGrew (2009), for further reading on contemporary theories of intelligence.

A disadvantage of the Wechsler scales in general (and most other current intelligence test) is that they do not cover the complete CHC taxonomy. Only five broad abilities (crystallized knowledge, visual processing, short-term memory, processing speed and fluid reasoning) are measured in both the WAIS-III (Alfonso, Flanagan & Radwan, 2005) and WAIS-IV (Grégoire, 2013; Weiss, Keith, Zhu & Chen, 2013a). Furthermore, multiple subtests can show loadings on the same ability, and specific abilities may not be completely covered by the subtests. In other words, CHC provides a rather complete structure of human intelligence, but it is challenging to develop tasks which are pure measures of those abilities. Describing and developing new intelligence tests within the CHC taxonomy would be advisable. The WAIS-III does not fit the theory well, and therefore describing the test only in terms of the CHC does not necessarily contribute to clinical evaluation of patients.

Another persisting criticism of the WAIS-III is that it disproportionately assesses Gc, in comparison to Gf (Blair, 2006, Duncan et al., 1995). Duncan et al. (1995) found unchanged WAIS-IQs in patients with frontal-lobe damage, while performance on a Gf task (Cattell's Culture Fair Task) was significantly impaired. Looking at the nature and location of the lesions, they also concluded that Gf was in fact a reflection of executive functioning (EF). EF can be defined as abilities which enable us to produce independent, purposive, self-directed and self-serving behavior (Lezak, Howieson, Bigler & Tranel, 2012). This includes (mental) adaptivity and flexibility, planning and problem solving capacities as well as (social) decision making skills.

Many studies suggest an extensive overlap between Gf and EF (Ardilla, 1999; Duncan et al., 1995; Duncan, Schramm, Thompson & Dumontheil, 2012; Van Aken, Kessels, Wingbermühle, van der Veld & Egger, 2015a; Roca et al., 2010), given the fact that both are related to effective performance in complex or novel situations, as well as frontal lobe functioning. For instance, all WAIS-III subtests added to enhance the measurement of Gf (Matrix Reasoning, Symbol Search, Letter-Number Sequencing) are related to EF performance (McGurk et al., 2000; Oosterman & Scherder, 2006; Sweet et al., 2005). This was studied using imaging techniques in both clinical and healthy samples. The subtest Digit Symbol Coding is also related to EF (Davis & Pierson, 2012). At the level of index scores, research showed affected PSI, WMI and, to a more limited extent, also POI in patients with brain injury and EF dysfunctioning (Ferry et al., 2004; Fisher et al., 2000; Hawkins, 1998; Martin, Donders & Thompson, 2000; Taylor & Heaton, 2001, Van der Heijden & Donders, 2003b). This raises the question to what extent EF is incorporated and distributed in the factor structure of the WAIS-III.

Both Wood & Liossi (2007) and Davis, Pierson & Holmes Finch (2011) examined the relation between intelligence and EF using the WAIS-III in a brain injured and healthy sample, respectively. Both studies demonstrated that EF is to some extent comparable to parts of global intelligence (*g*) measured by the FSIQ of the WAIS-III,

next to unique variance which seems to reflect more specific executive requirements ('*e*'; see Wood & Liossi, 2007). In 2001, Kaufman, Lichtenberger & McLean suggested a three-factor model solution for the WAIS-III, including verbal comprehension (factor 1), perceptual organization (factor 2) and a third factor labeled EF. For determining this third factor, they assumed that EF and working memory were interrelated, since the third factor was based on high loadings of Digit Symbol Coding and Letter Number Sequencing. They interpreted the third factor as being a blend of working memory and processing speed.

The aim of the present study is to gain insight in the degree to which EF is included in the WAIS-III in a heterogeneous sample consisting of both psychiatric patients and healthy volunteers. All 13 subtests of the WAIS-III are administered, except for Object Assembly given the poor low reliability of this subtest. To assess the broad construct of EF, Dutch versions of multiple traditional EF tasks were included: The Wisconsin Card Sorting Test (WCST; Heaton, 1981), the Stroop Color Word Test (Stroop; Hammes, 1971) and the Behavioural Assessment of the Dysexecutive Syndrome (BADS; Wilson, Alderman, Burgess, Emslie & Evans, 1996; Krabbendam & Kalff, 1997).

Based on earlier factor analytical studies of the Wechsler scales, in compliance with the four index scores, we expect a four factor structure to best fit the data. Nevertheless, since little research is done with all 13 WAIS-III subtests combined with different measures of EF, models with 2 (according to the VIQ-PIQ dichotomy) and 3 (according to Kaufman et al., 2001) factors will also be evaluated. In line with Wood & Liossi (2007), we expect the executive tasks to load high on the factors analogous to the PIQ scale (consisting of the POI and PSI) of the WAIS-III. More specific, we expect the BADS to load high on the factor corresponding to the POI, the WCST to load on either of the factors representing the POI or WMI, and the Stroop on the factor comparable to the POI or PSI.

Method

Participants

Included were 258 participants, (mean age 33.0 ± 14.1, 64.3 % male), consisting of 24 community-dwelling volunteers (mean age 37.2 ± 15.7, 45.8 % male) and 234 inpatients and outpatients (mean age 32.55 ± 13.81, 66.2 % male) of the Dutch Vincent van Gogh Institute for Psychiatry. In accordance with the guidelines of the institutional review board, records were drawn from a large electronic database, containing test results of patients admitted in the period from 2005 to 2013. Data were obtained as part of the standard neuropsychological assessment. Exclusion criteria for healthy volunteers were use of narcotics or sedatives and a presence or history of alcohol abuse, psychiatric illness or neurological disease.

All participants were Dutch-speaking. Psychiatric patients had a Full Scale IQ score (FSIQ) between 61 and 131 (M = 95.9; SD = 13.6). Healthy volunteers had a FSIQ scores between 79 and 141 (M = 106.6; SD = 15.4). Patients were diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR, 2000). Diagnoses included major affective (including bipolar) disorders (22.6%), anxiety disorders (7.8%), substance related disorders (5.1%), psychotic disorders (4.7%), dementia and other cognitive disorders (1%), developmental disorders (32.9%), adjustment disorders (7.3%), other disorders (3%; mainly identity and relational problems), and no diagnosis on axis I (7.8%). In some patients (8.1%) the formal diagnosis was unknown. Comorbidity with personality disorders was diagnosed in 24.3% of the patients, or diagnosis on axis II was deferred (29%).

Materials

The Dutch version of the Wechsler Adult Intelligence Scale - third edition was administered (except for the subtest Object Assembly) according to standard procedures (Wechsler, 2000). Reliability statistics are comparable to those found in the US version with internal consistency coefficients ranging from .72 for the subtest Picture Arrangement to .93 for the subtest Vocabulary. The factor structure of the Dutch WAIS-III is similar to the US version, except for Arithmetic, which has high loadings on both the VCI, POI and WMI, instead of a specific high loading on the WMI (Van der Heijden, Van den Bos, Mol & Kessels, 2013; Van Ravenzwaaij & Van Hamel, 2006). Raw scores of all subtests were included in analyses.

The Dutch version of the BADS (Krabbendam & Kalff, 1997), WCST (Heaton, 1981, Heaton, Chelune, Talley, Kay & Curtiss, 1993) and the Stroop (Hammes, 1971) were included as a comprehensive reflection of different EF sub functions. The BADS contains six subtests (Rule shift cards, Action Program, Key Search, Temporal Judgment, Zoo Map

and Modified Six Elements), which measure planning, problem solving, set-shifting, monitoring behaviour and the use of strategy (Lezak et al., 2012; Wilson et al., 1996). The *overall profile score*, computed out of standard scores of each subtest (not corrected for age, gender or education), was included in analysis.

The WCST is a test of abstract reasoning, requiring mental flexibility (set-shifting), problem solving skills and working memory (Heaton et al., 1993). Subjects have to sort cards with symbols, varying in shape, color and number, and achieve categories predetermined by the examiner, who gives feedback after each sorted card (‘right’ or ‘wrong’). Rules of how to sort cards are changed without warning after ten correct placements of the cards, and the participants had to adapt their strategy to the new rule and change their responses. As an overall performance index (Lezak et al., 2012), the *total number of errors* was used in analyses.

The Stroop is a test of response inhibition and cognitive flexibility. Subjects have to read colors out loud on three cards. The first card shows the name and print of the color, card two shows only colors and the third card shows a color in words with an incongruent color print, in which the latter has to be read out loud. To measure the concept of response inhibition, the element of speed is eliminated by using the *interference score* (response time on card III divided by the average response time of card I + II) for analysis.

Analyses

Factor analysis was conducted using PASW Statistics (version 18). The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) was used to determine the appropriateness of the factor analysis (Field, 2009). Maximum likelihood procedures with promax rotation were used to examine the model fit of two, three and four factors. The chi-square statistic was used to determine the goodness of fit of all three factor solutions. Taking a conservative approach (Field, 2009), only factor loadings ≥ .35 were interpreted.

Results

Descriptive statistics of all tests are presented in Table 1. Intercorrelations of the WAIS-III and the measures of EF are shown in Table 2. The KMO score is .91, which can be considered excellent (Field, 2006).

Table 3 shows the goodness of fit indices of two, three and four factor models. Both 3 factor and 4 factor models fit the data. The model with four factors fit the data best, and therefore was considered to be the best fit for the current sample. Table 4 presents the factor loadings in the four factor structure.

All factor loadings of the WAIS-III subtests were comparable to the four factor indices of the WAIS-III (VCI, POI, PSI and WMI) except for Arithmetic, which loaded on the VCI instead of the WMI in the original WAIS-III manual (Wechsler, 1997b), and Letter-Number Sequencing, which loaded on both POI and WMI. All executive tasks had their highest loading on factor two, which is comparable to the POI of the WAIS-III.

The correlations between the four factors are shown in table 5. All factors are strongly correlated, except factor one (VCI) and factor three (PSI), which have a medium sized correlation (Cohen, 1992). Factor two (POI) and three (PSI) share the highest correlation of .64.

Table 1 Descriptive Statistics for the WAIS-III Subtests, BADS, WCST, and Stroop

Test	Patients (*n*=234) Mean	SD	Healthy volunteers (*n*=24) Mean	SD
WAIS-III Subtests				
Picture completion	20.3	3.3	21.0	2.5
Vocabulary	39.3	11.3	45.8	10.7
Digit Symbol-Coding	66.0	17.0	76.0	14.8
Similarities	24.7	4.7	27.3	4.1
Block Design	37.4	16.5	42.8	19.6
Arithmetic	12.1	4.7	14.9	3.3
Matrix Reasoning	17.9	5.1	20.1	5.0
Digit Span	14.7	3.4	16.8	4.1
Information	15.2	5.2	17.1	5.5
Picture arrangement	12.7	4.7	15.3	3.7
Comprehension	23.0	5.3	26.3	4.1
Symbol Search	31.6	8.1	35.5	8.2
Letter-Number Sequencing	9.9	3.0	11.5	2.3
BADS	18.4	3.1	18.7	2.5
Stroop	1.9	0.5	1.7	0.3
WCST	24.7	18.1	17.0	12.5

Note. WAIS-III = Wechsler Adult Intelligence Scale - Third Edition, BADS = Behavioural Assessment of the Dysexecutive Syndrome, WCST = Wisconsin Card Sorting Test

Table 2 Intercorrelations of the 13 WAIS-III Subtests and BADS, WCST, and Stroop

	PC	V	DSC	S	BD	A	MR	DS	I	PA	C	SS	LNS	BADS	Stroop	WCST
PC		.33*	.36*	.33*	.49*	.33*	.55*	.24*	.37*	.47*	.30*	.39*	.35*	.46*	-.26*	-.22*
V			.33*	.72*	.29*	.58*	.35*	.42*	.78*	.28*	.78*	.16*	.38*	.23*	-.10	-.15*
DSC				.39*	.49*	.38*	.49*	.44*	.32*	.41*	.26*	.66*	.35*	.32*	-.31*	-.35*
S					.39*	.55*	.47*	.48*	.64*	.38*	.62*	.25*	.39*	.32*	-.21*	-.22*
BD						.43*	.63*	.36*	.39*	.50*	.23*	.61*	.38*	.40*	-.35*	-.34*
A							.50*	.43*	.60*	.28*	.56*	.35*	.46*	.34*	-.22*	-.23*
MR								.39*	.41*	.51*	.29*	.50*	.47*	.44*	-.31*	-.34*
DS									.42*	.32*	.32*	.40*	.60*	.30*	-.29*	-.24*
I										.33*	.68*	.22*	.42*	.27*	-.19*	-.21*
PA											.23*	.43*	.37*	.41*	-.31*	-.32*
C												.12**	.32*	.23*	-.13**	-.20*
SS													.34*	.34*	-.37*	-.32*
LNS														.33*	-.30*	-.30*
BADS															-.23*	-.27*
Stroop																.18*

Note. PC = Picture Completion, V = Vocabulary, DSC = Digit-Symbol Coding, S = Similarities, BD = Block Design, A = Arithmetic, MR = Matrix Reasoning, DS = Digit Span, I = Information, PA = Picture Arrangement, C = Comprehension, SS = Symbol Search, LNS = Letter-Number Sequencing, BADS = Behavioural Assessment of the Dysexecutive Syndrome, WCST = Wisconsin Card Sorting Test.

* = Significant at p < .01

** = Significant at p < .05

Table 3 Goodness of Fit for 2, 3 and 4 Factor Models Based on the Maximum Likelihood Procedure

Factors	χ^2	df	p
2	191.33	89	.000
3	129.09	75	.000
4	68.38	62	.270

Note. df = Degrees of Freedom

Table 4 Rotated Four Factor Solution

	Factors			
Test	I	II	III	IV
Vocabulary	**1.02**	-.16	.02	-.01
Comprehension	**.92**	-.09	-.01	-.08
Information	**.80**	.08	-.02	.00
Similarities	**.65**	.15	-.05	.11
Arithmetic	**.51**	.20	.08	.05
Matrix Reasoning	.01	**.84**	-.05	-.01
Picture Completion	.10	**.72**	-.03	-.15
Picture Arrangement	-.01	**.67**	.00	-.00
Block Design	-.04	**.66**	.21	-.06
BADS	-.04	**.62**	-.07	.04
WCST	.03	**-.39**	-.06	-.03
Stroop	.12	**-.35**	-.13	-.12
Symbol Search	-.01	.05	**.99**	-.03
Digit Symbol-Coding	.12	.16	**.48**	.10
Digit Span	.00	-.10	.03	**1.00**
Letter-Number Sequencing	.05	.34	-.09	**.46**
Eigenvalue	6.68	2.06	1.00	.84
% of variance	41.73	12.89	6.23	5.24

Note. BADS =Behavioural Assessment of the Dysexecutive Syndrome, WCST = Wisconsin Card Sorting Test. All factor loadings ≥ 0.35 are highlighted in boldface.

Table 5. Correlations Between Factors

Factor	I	II	III	IV
I		.56	.27	.53
II			.64	.56
III				.45

Discussion

The goal of the present study was to gain insight in how executive functioning is incorporated in the WAIS-III, examined in a large heterogeneous group of both psychiatric patients and healthy volunteers. Exploratory factor analysis using maximum likelihood procedures was conducted to examine the model fit of two, three and four factor models. Both three and four factor models fit the data. The three factor model was almost identical to the model found by Kaufman et al. (2001), and consisted of a combined POI/PSI factor, and a second and third factor which were completely comparable to the VCI and WMI of the WAIS-III, except for Arithmetic. The four factor model was almost completely identical to the four factor indices of the WAIS-III. The BADS, WCST and Stroop loaded all on factor two, corresponding to the POI of the WAIS-III. This model fitted the data best, and therefore was selected for interpretation.

Contrary to expectation, Arithmetic rather loads on the VCI than on the WMI in the American WAIS-III manual. This is also in contrast with the Dutch manual of the WAIS-III, in which Arithmetic loads on all index factors except PSI. This emphasizes the unstable structure of this subtest (Grégoire, 2013; Ravenzwaaij & Van Hamel, 2005; Van der Heijden, Van den Bos, Mol & Kessels, 2013). Furthermore, the intercorrelation of the VCI and PSI was low (.27), which differs from the intercorrelation of the two index scores in the Dutch WAIS-III manual (.50). According to Hawkins (1998), and Taylor & Heaton (2001), PSI can be considered the most sensitive factor for various clinical disorders and VCI the least. This is probably due to the fact that VCI is a measure of 'hold' tasks, which means performance stays relatively uninfluenced by brain disease or impairment, and this may explain the low correlation between VCI and PSI in the current sample.

Results support the hypothesis that components of *g* can be measured by EF tasks (Duncan, 1995, Wood & Liossi, 2007). The BADS, WCST and Stroop have loadings of respectively .62, -.39 and -.35 on the POI, implicating that this index accounts at least for some part of variance in EF performance. These results are similar to results found by Wood & Liossi (2007), who concluded that performance on all neuropsychological tests of executive function correlated with the WAIS-III FSIQ and PIQ scores. The partial reflection of EF performance in the POI is in line with the upcoming evidence of a great overlap between EF and Gf, since (subtests of) the PIQ scale are often associated with Gf (Duncan, 2010; Roca et al., 2010; Van Aken et al., 2015a).

The explained variance by the POI accounts for 12.89 % in the model, compared to 41.73 % by the VCI. Given the importance of executive functions in neuropsychological evaluation of patients as well as their relation to intellectual (dis)abilities, these results contribute to the already persisting criticism that the WAIS-III is mainly a test of crystallized intelligence (Blair, 2006). Since *g* for the greater part can be explained by Gf (Duncan et al., 1995), this should translate to the distribution of more Gf and EF subtests, instead of an overload of Gc subtests. In terms of CHC, this is in agreement with the suggestion of Ward, Bergman & Hebert (2012) and Grégoire (2013), who propose a more hierarchic structure in the description of CHC abilities. More specific, they state that fluid reasoning should go upwards in the hierarchy, given its influence on *g* and its impact on overall (cognitive) functioning.

As to the relation between Gf and EF, it is suggested that they co-exist through a general Gf factor next to more specific EF sub processes like set shifting, inhibition or processing speed (Duncan et al., 2008; Miyake et al., 2000; Wood & Liossi, 2007). In Duncan (2010), Gf is hypothesized as being a reflection of the efficiency in which complex behavior (consisting of different EF processes) is set up. The more complex or novel the task demands, the more interference of Gf is required. Therefore, including task complexity as an important variable in intelligence studies, more insight in the relation of intelligence and EF will be gained. Nevertheless, contemporary research should keep in mind that in the current, but nearly all factor analytic studies on intelligence, covariance between both subtests and factors exists, just as they will interact in daily life. Therefore, interpreting cognitive disabilities within the context of the individual is essential for both assessment and treatment of cognitive disorders.

A limitation of the current study might be the use of explorative factor analysis, which makes it less unequivocal to compare with many other studies on intelligence, which tend to utilize confirmative methods (Bowden, 2013). This decision was made based on the fact that little (confirmative) research is done combining the WAIS-III subtest and additional (EF) tasks. Nevertheless, using the maximum likelihood procedure, results could still be interpreted within the four-factor structure known from the WAIS-III. In the future, a consideration would be to include executive tasks with higher reliability and validity statistics, using more robust statistical analyses to support the theory. Moreover, it is evident that, if available, the use of the WAIS-IV in future research is preferred given the more prominent influence of (especially) Gf in the development of the subtests.

Conclusion

The current study gives more insight into the distribution of EF in the WAIS-III. Limited performance on the POI gives direction to further examination of different EF aspects, which in turn could account for disharmonic distributions in intellectual abilities. Although current models of intelligence and EF tend to describe process pure abilities to find an overall theory of abilities, the assessment and treatment of cognitive disorders requires a multifaceted and integrated view in which cognitive disorders can be understood through the interaction of an individual and the environment.

Chapter 3

Representation of the Cattell-Horn-Carroll Theory of Cognitive Abilities in the Factor Structure of the Dutch-language Version of the WAIS-IV

Van Aken, L., Van der Heijden, P.T., Van der Veld, W.M., Hermans, L., Kessels, R.P.C., & Egger, J.I.M. (2015). Representation of the Cattell-Horn-Carroll Theory of Cognitive Abilities in the Factor Structure of the Dutch-language Version of the WAIS-IV. *Assessment,* 1-9. doi:1073191115607973.

Abstract

The Cattell-Horn-Carroll (CHC) theory of cognitive abilities has been guiding in the revision of the Wechsler Adult Intelligence Scale – Fourth edition (WAIS-IV). Especially the measurement of fluid reasoning (Gf) is improved. A total of five CHC-abilities are included in the WAIS-IV subtests.
Using confirmatory factor analysis, a five-factor model based on these CHC abilities is evaluated and compared with the four index scores in the Dutch-language version of the WAIS-IV. Both models demonstrate moderate fit, preference is given to the five-factor CHC model both on statistical and theoretical grounds.
Evaluation of the WAIS-IV according to CHC terminology enhances uniformity, and can be important when interpreting possible sources of index discrepancies. To optimally assemblage CHC and WAIS-IV, more knowledge of the interaction of abilities is needed. This can be done by incorporating intelligence testing in neuropsychological assessment. Using this functional approach contributes to a better understanding of an individual's cognitive profile.

Introduction

The Wechsler Intelligence Scales are the most frequently used measures of intelligence worldwide (Lichtenberger & Kaufman, 2009). In 2008, the revised Wechsler Adult Intelligence Scale – Fourth edition was introduced (Wechsler, 2008). Goals of the revision were to enhance theoretical underpinnings, psychometric quality, and to improve user friendliness as well as clinical utility (Wechsler, 2008). The greatest modification in the fourth edition is the elimination of the verbal and performance IQ scales (VIQ and PIQ). Lacking theoretical ground and structural validity, this dichotomy is now replaced by the four indices Verbal Comprehension (VCI; three subtests and one supplemental subtests to measure verbal reasoning abilities which require comprehension and conceptualization), Perceptual Reasoning (PRI; three subtests and two supplemental subtests as a measurement of non-verbal reasoning and perceptual organization), Working Memory (WMI; two subtests and one supplemental subtest measuring attention, concentration and working memory), and Processing Speed (PSI; two subtests and one supplemental subtest to measure speed of mental and graphic-motor processing), from which a composite IQ score can be made up (i.e. Full Scale IQ; FSIQ; Wechsler, 2008). This structure is in line with Wechsler's view of intelligence as being a global entity representing overall functioning in daily life, which can also be made up of different cognitive domains (Grégoire, 2013).

The subtests of the WAIS-IV are developed from the latest theoretical perspectives on intelligence and cognitive neuroscience. Especially fluid reasoning (Gf), working memory and processing speed have been of particular importance in developing and adjusting subtests (Lichtenberger & Kaufman, 2009). For example, to strengthen measurement of Gf, the new subtests Figure Weights and Visual Puzzles are introduced as measures of PRI. Number sorting is added to Digit Span, and Cancellation is added as an optional measure of PSI (Wechsler, 2008).

Recently, researchers have examined the WAIS-IV factor structure in terms of the Cattell-Horn-Carroll theory of cognitive abilities (CHC theory; McGrew, 2009). This taxonomy is a combination of Horn & Cattell's theory of fluid and crystallized intelligence (Gf-Gc theory; 1966), and Carroll's three stratum theory (Carroll, 1993). CHC theory consists of three strata: general intelligence or *g,* also known as Spearman's *g* (Spearman, 1927), about ten to sixteen broad cognitive abilities, and over one hundred narrow abilities (McGrew, 2009, Schneider & McGrew, 2012). The abilities described in this theory have been of great importance in the development of intelligence batteries over the past three decades (e.g. the Wechsler scales, Woodcock-Johnson editions, Kaufman scales, et cetera; Keith & Reynolds, 2010), and the aggregated CHC theory has become more and more guiding in the development of subtests (Alfonso, Flanagan & Radwan, 2005; Keith & Reynolds, 2010). The CHC theory proposes the inclusion of five broad cognitive abilities in the WAIS-IV (Alfonso et al., 2005); crystallized knowledge (Gc), fluid reasoning (Gf), short-term memory (Gsm), visual processing (Gv), and processing speed (Gs).

Although Wechsler (2008) endorsed a higher-order four factor structure in the technical manual of the WAIS-IV, several studies report psychometric support for alternative, theory driven (five-factor) structures according to CHC in the Wechsler scales (Benson, Hulac, & Kranzler, 2010; Golay, Reverte, Rossier, Favez, & Lecerf, 2012; Niileksela, Reynolds, & Kaufman, 2013; Ward, Bergman, & Hebert, 2012; Weiss, Keith, Zhu, & Chen, 2013a). However, it has been discussed that a bi-factor model (also known as a nested factor model or direct hierarchical model, allowing subtests to load directly onto *g*) is superior to the division into index scores, and that interpretation of the WAIS-IV should be mainly at the level of *g* (Canivez & Watkins, 2010; Gignac & Watkins, 2013; Niileksela et al., 2013).

Difficulties of analyzing WAIS-IV performance in terms of CHC abilities are (a) that subtests are not designed as dedicated measures of isolated CHC abilities, and (b) that they do not cover all CHC abilities. In addition, (c) it is unknown if subtests measure the entire scope/breath of such an ability (Grégoire, 2013). This is a consequence of the multi-factorial nature of the WAIS-IV subtests that were initially selected for their clinical usefulness. Therefore, direct comparison of index scores and CHC abilities is not possible and interpretation according to CHC is complex. However, the incremental value of explaining the WAIS-IV in terms of CHC is that performance can be understood within a theoretical framework that uses a common nomenclature among researchers. Ongoing study on the CHC taxonomy will further improve and extend the current model.

Benson and colleagues (2010) propose a five-factor structure for the 15 subtests of the WAIS-IV which leads to the subdivision of PRI into Gf (defined by Matrix Reasoning, Figure Weights, and Arithmetic) and Gv (defined by Block Design, Visual Puzzles, and Picture Completion). Arithmetic demonstrates cross loadings on both Gf and Gsm (identical to the WMI), but is primarily a measure of Gf in this five-factor-structure. Furthermore, to improve the model, they suggest an underlying narrow ability Quantitative Reasoning (QR) under Gf, defined by Figure Weights and Arithmetic. Similarly, a study of Weiss et al. (2013a) on the WAIS-IV normative sample results in an almost identical five factor model. Furthermore, they demonstrated factor invariance across the clinical and non-clinical normative samples of the WAIS-IV, supporting the robustness of this model across different samples.

From a theoretical perspective, the five-factor based model is recommended since the interpretation of PRI does not follow CHC structure. The subdivision of PRI into Gf and Gv conforming CHC exemplifies a better fit (Benson et al., 2010; Weiss et al., 2013a). Niileksela et al. (2013) demonstrate that the five-factor model as proposed by Benson et al (2010) and Weiss et al (2013a) fit the WAIS-IV data of a sample of elderly; although Gv and Gf are highly correlated (.96), they are distinct. Furthermore, they state that the five-factor model has fewer misspecifications than the four-factor-model. Therefore, they conclude that dividing PRI into Gv and Gf is valuable, but the measurement of the CHC abilities is limited, since only a few of the subtests of these factors are suitable for elderly people. Furthermore, Weiss et al. (2013a) state that this model has more possibilities for interpretation in disharmonic profiles, since it offers explanation of profile inconsistencies by so called secondary abilities, based on the found cross loadings in the model. Next to the primary interpretation following the four-factor structure, they provide an alternative secondary (five-factor structure) interpretation when discrepant scores between indices are found.

Although Weiss and colleagues (2013a) provide psychometric support for a five-factor model, the possible benefits of a fifth factor over the four index scores in terms of clinical usefulness and interpretation has not lead to consensus yet. The primary and secondary interpretive models as proposed by Weiss et al. (2013a) did receive some criticism, since this decision (depending on profile discrepancies) is based on subtest scatter, and leads to poor reliability, validity, and diagnostic utility at the individual level (Canivez & Kush, 2013; Schwartz, 2013). Other criticism has focused on the separation of PRI in Gv and Gf. Especially the measure of Gf is weak, since it contains only three subtests which also have cross-loadings on other factors. This led Weiss et al. (2013a) to include the intermediary factor QR under Gf, which undermines the interpretation of the WAIS-IV in CHC terms (Canivez & Kush, 2013). Furthermore, interpretation of Gf is rather complex, partially due to the fact that extremely high loadings of Gf on *g* keep recurring in analyses, raising the question whether *g* and Gf are identical (Benson et al., 2010; Niileksela et al., 2013; Schneider & McGrew, 2012; Weiss et al., 2013a). As Grégoire (2013) states, interpretation of Gf as a separate factor in the WAIS-IV based on three subtests is not easier than interpretation of the current PRI. Altogether, clinicians should bear in mind that psychometric strength (and the risk of over-factoring; Frazier & Youngstrom, 2007) is not identical to clinical usefulness.

The current study aims to examine whether a five-factor structure according to CHC abilities is tenable in the Dutch-language version of the WAIS-IV (WAIS-IV-NL) in a psychiatric sample. Next to the four-factor structure according to the index scores of the WAIS-IV (Wechsler, 2008), the final five-factor CHC model as described by Weiss et. al. (2013a) in the US normative sample is investigated in the present study. Model fit will be examined using confirmatory factor analysis. Benson et al. (2010) and Weiss et al. (2013a), found that both four- and five-factor models demonstrate acceptable statistical fit. Similar results are expected in our clinical sample. In the case the five-factor model does show adequate fit, preference is given to this model, since it justifies the use of theoretical CHC underpinnings, which will enhance clinical interpretation of test performance. Evaluation of factor intercorrelations, subtest loadings on factors and internal consistency of factors will further contribute to the discussion on the theoretical relevance clinical usefulness of both models.

Method

Participants

Included were 233 psychiatric patients (M_{age}=35.07, *SD*=14.84, 52% male), consisting of 74 in- and outpatients from the Vincent van Gogh Institute for Psychiatry in Venray, and 159 patients from the Reinier van Arkel Mental Health Institute in Den Bosch, both located in the Netherlands. Psychological assessment (and thus WAIS-IV administration) is not a standard procedure in the admission at these institutes, so only the most complex cases with multiple DSM-IV diagnoses (including comorbidity with personality disorders) were referred for assessment, mostly being psychiatric patients with a history of school failure, cognitive complaints and/or impaired cognitive abilities. In accordance with the guidelines of the institutional review board, records were drawn from a large electronic database. For data analysis, patient identities were concealed. The majority of in- and outpatients received medical treatment to relieve symptoms of mental illness. The average FSIQ in the overall clinical sample was 87.1 (*SD* = 17.0), the average VCI was 90.6 (SD=16.8), the average PRI was 90.3 (SD=18.2), the average WMI was 88.0 (SD=16.9), and the average PSI was 86.8 (SD=17.7).

Materials

The WAIS-IV-NL was used (Wechsler, 2012a, 2012b). The test contains 10 core subtests (Block Design, Similarities, Digit Span, Matrix Reasoning, Vocabulary, Arithmetic, Symbol Search, Visual Puzzles, Information and Coding) and five additional subtests (Letter-Number Sequencing, Figure Weights, Comprehension, Cancellation, and Picture Completion). According to the WAIS-IV-NL technical manual (Wechsler, 2012a), split half reliability for the subtests vary between .75 and .93 in the complete sample. The index scores VCI, PRI, WMI, PSI and the FSIQ have split half reliabilities of respectively .96, .93, .92 and .88 and .97 (Wechsler, 2012a).

The construction of the subtests of the WAIS-IV-NL is similar to the WAIS-IV US edition, as described in the Dutch technical manual (Wechsler, 2012a). The WAIS-IV-NL follows a similar factor structure in comparison to the U.S. normative sample, with exception of the subtest Arithmetic, which correlates equally with the WMI, VCI, and PRI in the Dutch-language version (which was also the case in the Dutch-language version of the WAIS – third edition). Reliability coefficients for the Dutch language version of the WAIS-IV are comparable with those reported in the US manual (Wechsler, 2008; 2012a).

Procedure and analysis

The WAIS-IV was administered and scored according to the guidelines described in the manual (Wechsler, 2012b). Data were analyzed using LISREL 8.8 (Jöreskog & Sörbom, 2008). Confirmatory Factor Analysis (CFA) was conducted. Missing data

from the subtest CA were defined as missing values. The standardized subtest scores were used for analysis.

The four and five factor models as described by Weiss et al. (2013a) were evaluated. Multiple indices of model fit were used to analyze the proposed models; the chi-square (χ^2) statistic, the standardized root-mean-square residual (SRMR), the root mean square error of approximation (RMSEA) and 90% confidence intervals of this index, the comparative fit index (CFI) and the Akaike information criterion (AIC). RMSEA was used as an additional goodness-of-fit statistic next to χ^2, since it is corrected for complexity of models (Byrne, 2001; Thompson, 2000). An RMSEA <.08 is considered an acceptable to good fit (Schreiber, Nora, Stage, Barlow, & King, 2006). SRMR represents the average value across all standardized residuals. Values less than .08 are considered a relatively close fit (Schreiber et al., 2006). CFI indicates good fit when >.95 (Schreiber et al., 2006). To improve model fit, correlated errors were added to the model through evaluation of the modification indices from the initial output, if there were reasons of content (theoretical justifications) to do so. Since the four and five factor models were not nested, no χ^2 difference tests could be performed. Therefore, the AIC was used for comparison of the models; smaller AIC values indicate better fit.

Results

Subtest intercorrelations are presented in table 1. Both initial four- and five-factor models did not demonstrate acceptable fit (initial model fit statistics in table 2). In the four-factor model, modification indices suggested adding an error covariance between Information and Vocabulary and Block Design and Visual Puzzles. Subsequently, modification indices suggested adding error covariance between two first-order factors WMI and PSI, which are both known to be sensitive to neuropsychological deficits that may be present in our clinical sample (Kaufman, Lichtenberger & McLean, 2001; Kooij & Dek, 2012; Lichtenberger & Kaufman, 2009). Adding correlated errors between WMI and PSI resulted in moderate model fit, without changing the estimates of factor loadings (and thereby preserving factor structure; see final four-factor model, table 2).

Regarding the five-factor model, similar adjustments were made by introducing correlated errors between Information and Vocabulary and Gsm and Gs, based on the modification indices of the initial model. Furthermore, the five-factor model according to Weiss et al. (2013a) could not be replicated due to estimation problems for the factor Quantitative Reasoning. As an equivalent alternative, correlated errors between Figure Weights and Arithmetic were added. See table 2 for a summary of all fit statistics.

Table 1 Subtest Intercorrelations

	BD	Si	DS	MR	Vo	Ar	SS	VP	In	Co	LNS	FW	C	Ca	PC
BD	1														
Si	.53	1													
DS	.36	.48	1												
MR	.62	.61	.46	1											
Vo	.42	.73	.41	.52	1										
Ar	.48	.58	.56	.60	.64	1									
SS	.53	.42	.49	.47	.28	.45	1								
VP	.68	.56	.40	.63	.50	.49	.48	1							
In	.48	.71	.41	.54	.81	.65	.29	.51	1						
Co	.51	.46	.54	.49	.38	.45	.67	.46	.35	1					
LNS	.40	.48	.67	.42	.48	.62	.54	.39	.43	.58	1				
FW	.61	.63	.49	.63	.61	.67	.49	.60	.60	.46	.55	1			
C	.48	.75	.42	.58	.77	.62	.40	.49	.67	.50	.53	.63	1		
Ca	.52	.44	.38	.46	.30	.33	.52	.51	.26	.63	.46	.40	.31	1	
PC	.41	.41	.35	.40	.28	.34	.40	.47	.31	.39	.34	.40	.41	.29	1

Note. BD=Block Design, Si=Similarities, DS=Digit Span, MR=Matrix Reasoning, Vo=Vocabulary, Ar=Arithmetic, SS=Symbol Search, VP=Visual Puzzles, In=Information, Co=Coding, LNS=Letter-Number Sequencing, FW=Figure Weights, C=Comprehension, Ca=Cancellation, PC=Picture Completion.

Table 2 Fit statistics of the initial and final four and five factor models

4 factor model	χ^2	df	*p*	RMSEA	90% CI RMSEA	SRMR	CFI	AIC
initial	276.37	84	.00	.101	.088 - .110	.063	.97	348.37
final	225.09	81	.00	.089	.075 - .100	.051	.98	303.09
5 factor model	χ^2	(df)	*p*	RMSEA	90% CI RMSEA	SRMR	CFI	AIC
initial	254.30	82	.00	.097	.083 - .110	.064	.97	330.30
final	201.88	79	.00	.083	.069 - .098	.057	.98	283.88

Note. df = degrees of freedom; CI = confidence interval; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index; AIC = Akaike information criterion.

3

The final four-factor structure according to the WAIS-IV index scores is presented in Figure 1. As described, starting point of the current model was the final model described by Weiss et al. (2013a), therefore, a cross loading with Arithmetic is present. The final five-factor CHC model is presented in Figure 2. Estimates of factor loadings in this model are almost equal to the Weiss et al (2013a) five factor model.

The decline of AIC in the five-factor model did demonstrate improvement in model fit, favoring this model over the four-factor model. In the latter, the correlated error between Visual Puzzles and Block Design already implies the existence of a Gv factor. In the five-factor model, the definition of Gf is fragile, given the cross loadings of Matrix Reasoning and Figure Weights on Gv. Looking at the factor structure of both models (Table 3 and 4), factor intercorrelations are generally comparable in both models. In the five-factor model, factors tend to correlate higher with Gf than Gv. Finally, both models show very high loadings of Gf and PRI on *g,* a result also found in previews studies (Benson et al., 2010, Niileksela et al., 2013, Weiss et al., 2013a).

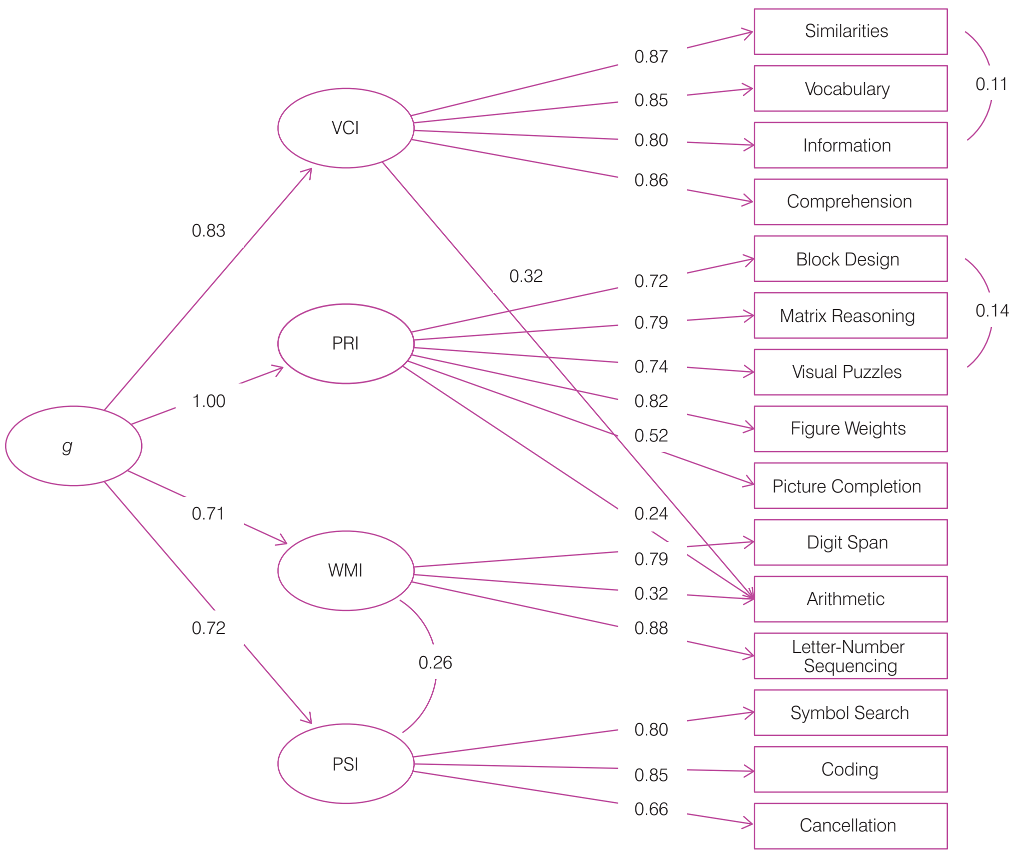


Figure 1 Final four-factor model

Note. g = general intelligence; VCI = Verbal Comprehension Index; PRI = Perceptual Reasoning Index; WMI = Working Memory Index; PSI = Processing Speed Index.

Table 3 Intercorrelations between factors of the four-factor-model

	VCI	PRI	WMI	PSI
VCI	1.00			
PRI	.83	1.00		
WMI	.59	.71	1.00	
PSI	.60	.72	.77	1.00

Note. VCI = Verbal Comprehension Index; PRI = Perceptual Reasoning Index; WMI = Working Memory Index; PSI = Processing Speed Index.

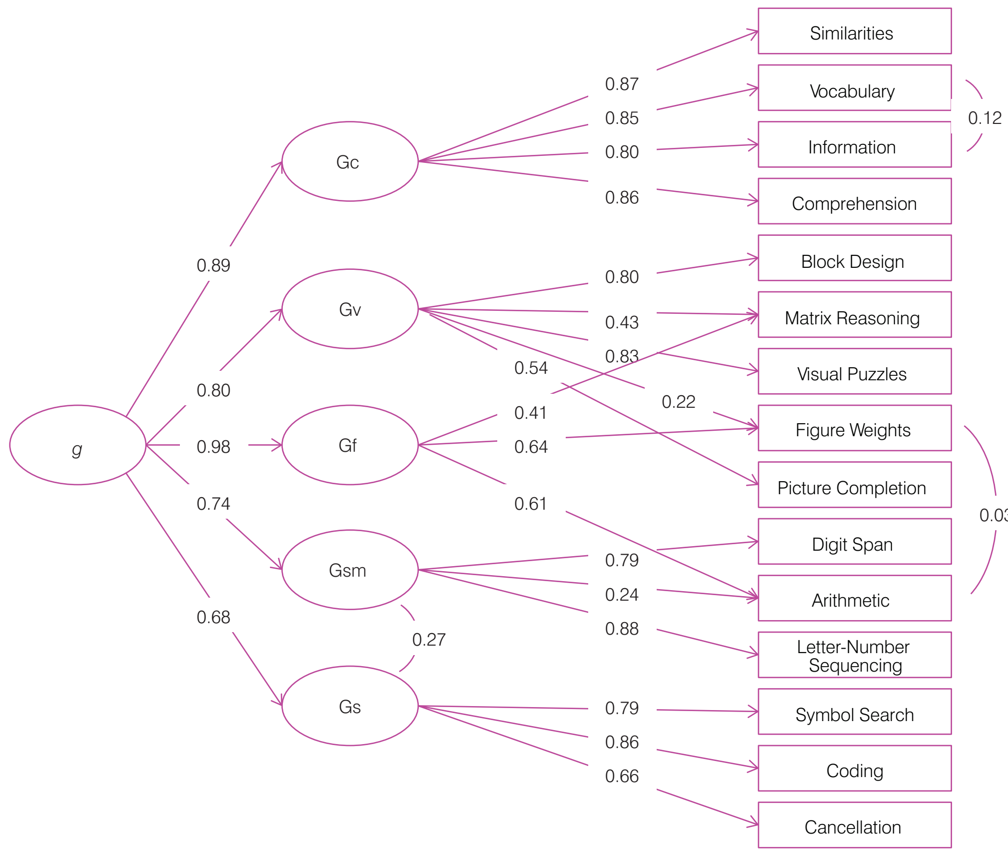


Figure 2 Final five-factor model

Note. g = general intelligence; Gc = crystallized intelligence; Gv = visual processing; Gf = fluid intelligence/reasoning; Gsm = short-term memory; Gs = processing speed.

Table 4 Intercorrelations between factors of the five-factor-model

	Gc	Gv	Gf	Gsm	Gs
Gc	1.00				
Gv	.71	1.00			
Gf	.87	.78	1.00		
Gsm	.66	.59	.73	1.00	
Gs	.61	.55	.67	.78	1.00

Note. Gc = crystallized intelligence; Gv = visual processing; Gf = fluid intelligence/reasoning; Gsm = short-term memory; Gs = processing speed.

Discussion

The main purpose of the present study was to look at the statistical fit and clinical usefulness of the four-factor model proposed in the WAIS-IV manual and a five-factor model derived from CHC theory. Both models demonstrate moderate fit. Although no difference test could be calculated since the models are not nested, psychometric evidence based on the AIC is in favor of the five-factor CHC model. Also, theoretical underpinnings support the preference of this model. Moreover, the increase in model fit of the four-factor model after adding error covariance between Block Design and Visual Puzzles, also suggests the existence of an underlying Gv factor. Even though the Gf-Gv correlation in the five factor CHC model is high, their different relations to the other factors and to *g* implicate they are in fact separate entities. The division of PRI into Gv and Gf therefore seems legitimate. Furthermore, current factor loadings in this model are almost equal to the factor structure of Weiss et al. (2013a) executed in the U.S. normative sample of the WAIS-IV.

The measurement of Gf is improved in the WAIS-IV, however, the amount of Gf in the whole test stays restricted and Gf subtests show cross loadings with other factors. Moreover, correlations between Gf and other factors are substantial. These results are comparable to results of Weiss et al. (2013a), who explained these cross-loadings in terms of secondary abilities and interrelations of cognitive abilities. Current results show that Gf is almost indistinguishable from *g,* which is also in line with previous research (Benson et al., 2010; Niileksela et al., 2013; Schneider & McGrew, 2012; Weiss et al., 2013a). Some authors argue that this equivalence is a mere statistical artifact (Golay et al., 2012), while others state Gf is isomorphic to *g* on a content level (Schweizer, Troche & Rammsayer, 2011). The discussion whether the two are identical or not, remains unresolved so far (Schneider & McGrew, 2012). Both the cross loadings and equivalence with *g* are following the features of Gf. Pure measurement of Gf is not feasible without addressing other abilities. It includes non-specific learning capacities which are needed for many types of (complex) behavior, therefore the observed overlap with both general and specific cognitive abilities is not surprising. Gf (next to Gc) has special properties apart from the other broad abilities and should be considered going upwards in CHC hierarchy. By all means it should not be interpreted as equal to other broad abilities. As Weiss et al. (2013b) state: ‘Perhaps Gf, when conceptualized as an integrative ability, is the ecological *g* that has eluded researchers for more than a century’.

The current study uses data collected in the clinical field. This is of high importance, since the WAIS-IV is one of the most frequently used intelligence test in clinical practice (Evers et al., 2012; Lichtenberger & Kaufman, 2009). Clinical data may demonstrate different properties than normative data. For instance, the mean FSIQ of the current sample is below average. Furthermore, subtest and factor intercorrelations of this sample are evidently higher than those described in the Dutch normative sample (Wechsler, 2012a). Moreover, the need for error covariance between latent constructs (WMI and PSI or Gsm and Gs) is unusual, and suggests a relation that goes beyond *g.* Altogether, the data result in moderate fit, meaning other variables have influence on performance too. The cause of this unknown variance lies, at least partially, in the psychiatric nature of the sample. Psychopathology, medical treatment and neuropsychological deficits may affect test performance (and especially influences scores on the WMI and PSI; Kooij & Dek, 2012; Lichtenberger & Kaufman, 2009). Unfortunately, we were not able to examine these effects in more detail, because precise diagnostic information (in terms of DSM-IV diagnoses) is unknown. However, the equality of the current five-factor CHC model with the results found by Weiss et al. (2013a), endorse our conclusions based on the current results.

The clinical usefulness of the results has previously been subject to debate. One of the remarks contains the addition of correlated errors (or lower-order factors as the QR factor in the model by Weiss et al., 2013a) to the model. Although this does increase model fit (without harming the structure or scale of the estimated factor loadings), it does not contribute to more straightforward interpretation (Canivez & Kush, 2013). Others state that, regardless whether a four- or five-factor structure is used, interpretation of the first-order factors should be avoided and clinicians should only interpret at the level of the FSIQ as a translation of *g* (see Canivez & Watkins, 2010, Gignac, 2008, and Gignac & Watkins, 2013, for further reading on this topic). Although these arguments are grounded, the fact that correlated errors are needed to enhance model fit, demonstrates that orthogonal interpretation of factors is simply not possible. Furthermore, reporting only the FSIQ does not contribute to the understanding of psychiatric patients and their performance on tasks and cognitive (dys)functioning in daily life. In other words, clinicians should embrace the multi-factorial structure of the subtests (which was exactly Wechsler’s purpose when developing them) to see the development and process of (affected) performance, especially in a clinical sample where lower or skewed profiles are common. For instance, comparing profile patterns on both the four indices and five CHC factors can expose inconsistencies, which in turn can be a starting point for further supplementary neuropsychological assessment to clarify these findings.

Thus, instead of striving for process pure subtests, information about which different abilities contribute to successful performance on one subtest would be more helpful for profile interpretation, which for instance is done in the Differential Ability Scale-II (Schwartz, 2013). McFarland (2013) concluded that models in which subtest performance is determined by multiple (uncorrelated) factors instead of the other way around (interpreting subtests as one dimensional functions of multiple correlated factors) result in better fit to the data. In other words, it makes sense that some subtests require both working memory capacities, speed of processing and/or visual processing, and limitations on one of these factors will lead to impaired performance on several subtests. Adopting the multidimensional nature of subtests creates more insight in how different abilities interact and lead to adequate performance, and shows what goes wrong if specific abilities are impaired. The next step is to examine *external* validity of the WAIS-IV. Examining external validity through comparison with other (neuro) psychological instruments will enhance diagnostic utility. For instance, executive functions are highly correlated to some aspects of intelligence and mainly Gf (Duncan, 2013). Such comparisons can enhance predictive power and clinical validity.

In conclusion, results contribute to clinical understanding of the WAIS-IV. The five factor model is valid, and division of the PRI in to Gf and Gv enhances understanding of performance on the corresponding subtests. Unfortunately, the measurement of Gf in the WAIS-IV remains restricted. Clinical awareness must be raised regarding the fallacy of orthogonal interpretation of scores of the WAIS-IV; index scores are not pure measures of isolated (CHC) abilities. Taking the FSIQ as a reliable starting point of a person's general level of performance, index scores can be used for refinement in the description of individual cognitive strengths and weaknesses. This can only be done through clinical experience with psychopathology and knowledge of the additional value of the index score.

Chapter 4

Fluid Intelligence and Executive Functioning more alike than different?

Based on: Van Aken, L., Kessels, R.P.C., Wingbermühle, E., Van der Veld, W.M., & Egger, J.I.M. (2015). Fluid intelligence and executive functioning more alike than different. *Acta Neuropsychiatrica, 28,* 31-37. doi:10.1017/neu.2015.46.

Abstract

Objective Fluid intelligence (Gf) has been related to executive functioning (EF) by previous studies, and it is also known to be correlated with crystallized intelligence (Gc). The present study includes representative measures of Gf, Gc, and EF frequently used in clinical practice to examine this Gf-EF relation. It is hypothesized that the Gf-EF relation is stronger than the Gc-EF relation, and that working memory in particular (as a measure of EF) shows a high contribution to this relation.

Method Confirmatory factor analysis was performed on a mixed neuropsychiatric and non-clinical sample consisting of 188 participants, using the Kaufman Adolescent and Adult Intelligence Test, and three executive tasks of the Cambridge Neuropsychological Test Automated Battery, covering working memory, planning skills and set shifting.

Results The model fit the data well [$\chi^2(24)=35.25$, $p=.07$, RMSEA=.050]. A very high correlation between Gf and EF was found (.91), with working memory being the most profound indicator. A moderate to high correlation between Gc and EF was present. Current results are consistent with findings of a strong relation between Gf and working memory.

Conclusion Gf and EF are highly correlated. Gf dysfunction in neuropsychiatric patients warrants further EF examination and vice versa. It is discussed that results confirm the need to distinguish between specific versus general fluid/executive functioning, the latter being more involved when task complexity and novelty increase. This distinction can provide a more refined differential diagnosis and improve neuropsychiatric treatment indication.

Introduction

The distinction between fluid (Gf) and crystallized intelligence (Gc), first made by Horn & Cattell (1966), has proven to be useful in neuropsychological assessment (Kaufman, Lichtenberger, & Kaufman, 2003). Gf is the ability to solve novel problems by using reasoning, and Gc is a knowledge-based ability that depends on schooling and acculturation (Horn & Cattell, 1966, Kaufman, Kaufman & Plucker, 2013). Gf and Gc have different functional properties. For instance, fluid abilities tend to decline from the age of 20, whereas Gc stays relatively preserved during ageing. Moreover, Gf is sensitive to brain damage, while Gc typically shows minor impairment after brain lesions (Duncan, Burgess & Emslie, 1995; Friedman, Miyake, Corley, Young, DeFries & Hewitt, 2006; Williams, Myerson & Hale, 2008). Examining general intelligence *(g),* fluid tests consistently appear to be its best predictors (Roca et al., 2014).

Executive functioning (EF) is a complex concept and contains multiple cognitive processes that are responsible for controlling and regulating thoughts, emotions, and behaviour and enable us to adjust to new situations (Diamond, 2013; Lezak, Howieson & Loring, 2004; Miyake, Friedman, Emerson, Witzki & Howerter, 2000). Miyake and colleagues (2000) identified updating, inhibition, and shifting as separate building blocks of EF, which together are a prerequisite for complex behaviour or 'higher-level executive functions' (Diamond, 2013). On the contrary, the unitary nature of EF becomes apparent in, for instance, the supervisory attentional system (SAS) by Norman & Shallice (1986). Being a contention scheduling based monitoring program, SAS selects sets of actions competing for representation and would thus be responsible for executive control of complex, goal-oriented behaviour. In recent years, researchers seemed to agree upon the approach that EF can be conceptualized as a unitary construct as well as consisting of diverse functions (Friedman et al., 2006; McCabe, Roediger, McDaniel, Balota & Hambrick, 2010).

Duncan and colleagues (1995) demonstrated that Gf is sensitive to frontal lobe lesions, leading to the conclusion that Gf is in fact a reflection of EF. Evidence from functional imaging studies further corroborates this overlap between Gf and EF in patients with frontal lobe lesions (Barbey, Colom, Paul & Grafman, 2014; Roca et al., 2010; Woolgar et al., 2010), Parkinson's disease (Roca et al., 2013a), fronto-temporal dementia (Roca et al., 2013b) and schizophrenia (Roca et al., 2014). In subsequent years, an increasing number of studies addressed the Gf-EF relation (Friedman et al., 2006; McCabe et al., 2010; Miyake et al., 2000; Salthouse & Davis, 2006; Salthouse & Pink, 2008; Salthouse, Atkinson & Berish, 2003). In general, Gf seems to correlate high with working memory (WM), whereas other aspects of EF (inhibition, mental set-shifting) usually show less strong relations with Gf (Duncan, Schramm, Thompson & Dumontheil, 2012; Friedman et al., 2006; Miyake et al., 2000; Redick, Unsworth, Kelly & Engle, 2012; Salthouse et al., 2003; Salthouse & Davis, 2006; Unsworth &

Engle, 2006; Unterrainer et al., 2010). Recently, Diamond (2013) concluded from a review of the literature that Gf can be regarded as being completely synonymous to the higher-level executive abilities reasoning and problem-solving.

The Kaufman Adolescent and Adult Intelligence Test (KAIT) is specifically designed to measure Gc and Gf (Kaufman & Kaufman, 1993). Apart from the Gf-Gc theory (Horn & Cattell, 1966), Luria's neuropsychological theory of intelligence (Luria, 1980) as well as Piaget's developmental concept of the formal-operational stage (Piaget, 2008), gave theoretical direction to the construction of the KAIT (Kaufman et al., 2003). Three widely used executive tasks from the Cambridge Neuropsychological Test Automated Battery (CANTAB) will be used to assess EF. The tasks include planning capacity and novel problem solving, working memory, reasoning, mental flexibility, and impulse control (Robbins et al., 1998). Although this is not an exhaustive sample of EF, a wide range of studied EF constructs is included, making the CANTAB tasks representative measures of EF.

Aims of the study

The present study examines the Gf-EF relation using a latent variable approach in a mixed sample of neuropsychiatric patients and non-clinical participants. In addition, it examines the relation of both Gf and EF with Gc. The main hypothesis is that EF, Gf and Gc are intercorrelated. Based on earlier research in which the Gf-EF relationship has been demonstrated in different (psychiatric) samples, we expect a Gf-EF relation higher than the Gc-EF relation. Furthermore, a high contribution of working memory to this relation can be expected, reflected in higher loadings of those CANTAB tasks on EF that appeal to working memory.

Method

Participants

Included were 188 participants (mean age 39.5±15.5, 51.6% male, N=98). This group consisted of 50 healthy individuals and 138 in- and outpatients of a neuropsychiatric department of a Dutch psychiatric hospital. See table 1 for demographic variables.

In accordance with the Diagnostic and Statistical Manual for Mental Disorders-Fourth Edition criteria, patients were diagnosed with major affective (including bipolar) disorder (44%), anxiety disorders (17%), impulsivity related psychopathology (9%), psychotic disorders (4%), dementia and other cognitive disorders (4%), developmental disorders (15%) and no formal psychiatric diagnosis (7%), respectively. Comorbidity with personality disorders was diagnosed in 37% of the patients.

For data analysis, patient identities were concealed. Informed consent was obtained from all healthy volunteers. Participants did not receive any compensation for participation.

Table 1 Demographics of the Sample Population

	N	% Male	Age (years)		Total IQ	
			M	SD	M	SD
Total	188	51.6	39.5	15.5	93.4	17.9
Patients	138	56.5	41.6	15.1	88.0	15.9
Healthy participants	50	38.0	33.5	15.2	111.9	10.8

In accordance with the guidelines of the institutional review board, patient records were drawn from a large electronic database, containing test results of patients admitted in the period from May 2007 to December 2012. The majority of in- and outpatients received medical treatment to relieve symptoms of mental illness.

Materials

Kaufman Adolescent and Adult Intelligence Test

The KAIT is an intelligence test for individuals between 11 and 85 years old and consists of a core battery containing six subtests (three Gf-tasks and three Gc-tasks), from which a composite IQ score can be made up. Test-retest reliabilities are good; 0.80 for Crystallized-IQ, 0.84 for Fluid-IQ and 0.89 for Total-IQ (Kaufman et al., 2003; Dekker, Mulder & Dekker, 2005; Kaufman, 2000).

The three fluid subtests focus on the integration of modalities and the efficiency of learning (Kaufman et al., 2003). *Rebus learning* contains associative learning and visual sequencing and requires intact working memory. *Mystery codes* measures speed of planning. *Logical Steps* assesses syllogistic reasoning and mathematics. The fluid subtests have reliabilities (Cronbach's α) of 0.91, 0.81, and 0.78, respectively (Mulder, Dekker & Dekker, 2005).

The three crystallized subtests contain the abilities of verbal understanding, verbal expression and verbal-conceptual development. *Definitions* measures the ability to deduct semantic relations, *Auditory comprehension* features auditory sequencing, and *Double meanings* requires semantic flexibility (Mulder et al., 2005). The crystallized subtests have reliabilities (Cronbach's α) of 0.84, 0.84, and 0.81, respectively (Mulder et al., 2005).

Cambridge Neuropsychological Test Automated Battery

The CANTAB is an automated test battery which has proven its utility for empirical research and in clinical practice (Parsey & Schmitter-Edgecombe, 2013). For further psychometric details on CANTAB tasks and indices, see Lowe & Rabbit (1998).

Spatial Working Memory (SWM) is a self-ordered working memory task that also assesses heuristic strategy, measuring the person's ability to retain and manipulate spatial information in the presence of interfering stimuli. Using a process of elimination, tokens have to be found in boxes. The boxes gradually increase in number and the position and colours keep changing per trial so stereotyped strategies are discouraged. The *number of between errors* (searching tokens in boxes that have been opened before) reflects a person's spatial working memory capacity (Owen, Downes, Sahakian, Polkey & Robbins, 1990) and was therefore selected for analysis.

The *Stockings of Cambridge (SOC)* is a task of planning and spatial working memory and refers to the ability to organise, plan and execute goal-directed behaviour (Lezak et al., 2004; Robbins et al., 1998). It is a computerized version of the tower tasks. Two displays with both three coloured balls are shown (which look like balls held in stockings). The fixed arrangement of balls in the upper display should be copied by the participant in the lower display. The minimum number of moves to complete the trial is shown on the screen, which increases from two to five moves. The *number of trials completed in the minimum number of moves* is selected as a measure of planning.

Intra-Extra Dimensional Set Shift (IED) is a test of rule acquisition and reversal. It is a computerized analogue of the Wisconsin Card Sorting Test and features maintenance, shifting and flexibility of attention. Two dimensions are used in the test, colour-filled shapes and white lines. Through the process of feedback and rule change, an intradimensional (shapes remain the only relevant dimension) and ex-tradimensional (lines become the only relevant dimension) set shift must be made. When failing to complete one block (six consecutive correct responses) after 50 trials, the test terminates. The *extra dimensional set shift errors* (block 8) are used as a measure of shifting (Robbins et al., 1998). If the task was cancelled before arriving at block 8, participants were given 25 errors on this block, the number of errors made based on chance.

Procedure

KAIT administration (paper-and-pencil) was followed by the CANTAB (computerized). Instructions were given in accordance with the standard administration in the user manuals. Participants were tested individually in a quiet environment. Mean testing time was approximately three hours.

Statistical analysis

Using Fisher r-to-z transformation on the sum-scores of Gf, Gc, and EF, group differences between healthy participants and psychiatric patients were tested. Confirmatory factor analysis was performed using LISREL 8.80 (Jöreskog & Sörbom, 2008) on raw data (n=188). Consequently, LISREL uses the *full information maximum likelihood (FIML) estimator.* The FIML procedure in LISREL only produces the FIML χ^2statistic and the root mean square error of approximation (RMSEA); no other fit indices are provided. A three-factor model was investigated to test the hypothesis that Gf, Gc, and EF are correlated. The factor models were evaluated using both goodness-of-fit measures and standardized factor loadings. As a rule of thumb, RMSEA < 0.05 indicates good fit (Hu & Bentler, 1999) and standardized factor loadings should be > 0.4.

Results

Descriptive statistics for the nine measures of EF and intelligence are presented in Table 2. Correlations between all measures are shown in Table 3. No group differences in correlations are found for the sum-scores of Gf, Gc, and EF between the psychiatric patients and healthy participants (z-scores all < 1.26, p-values all $> .20$). Some values of skewness and kurtosis are significant. However, multiple studies have shown that the *maximum likelihood* estimator is robust, under general conditions, against deviations from normality (Anderson & Amemiya, 1988; Satorra, 1992; Satorra & Bentler, 1990). Although the IED distribution approaches bimodality, this task is nevertheless included because it is necessary to examine the full scope of EF.

Table 2 Descriptive Statistics for the KAIT Subtests and CANTAB Indices (N=188)

Task	Range	Mean (SD)	Skewness	Kurtosis
KAIT				
Rebus Learning	19 to 99	69.95 (18.26)	-.41*	-.58
Logical Steps	0 to 16	7.81 (4.27)	.48*	-1.07*
Mystery Codes	4 to 34	18.65 (7.24)	.14	-.75*
Definitions	4 to 27	20.03 (4.20)	-1.22*	1.38*
Auditory Comprehension	1 to 18	9.74 (4.37)	-.08	-1.09*
Double Meanings	0 to 28	13.59 (5.55)	-.01	-.34
CANTAB				
Intra-Extra Dimensional Set Shift	0 to 32	9.04 (10.07)	1.08*	-.53
Stockings of Cambridge	0 to 12	8.50 (2.07)	-.65*	.97*
Spatial Working Memory	0 to 94	28.79 (22.08)	.60*	-.33

Note. KAIT: Kaufman Adolescent and Adult Intelligence Test, CANTAB: Cambridge Neuropsychological Test Automated Battery
* $p < .05$

Table 3 Intercorrelations of the KAIT subtests and CANTAB indices

	KAIT						CANTAB		
	Rebus Learning	Logical Steps	Mystery Codes	Definitions	Auditory Comprehension	Double Meanings	IED	SOC	SWM
KAIT									
Rebus learning	1								
Logical steps	.60**	1							
Mystery codes	.70**	.71**	1						
Definitions	.44**	.40**	.46**	1					
Auditory comprehension	.55**	.58**	.60**	.61**	1				
Double meanings	.49**	.51**	.54**	.63**	.54**	1			
CANTAB									
IED	-.25**	-.29**	-.31**	-.15*	-.26**	-.26**	1		
SOC	.42**	.44**	.43**	.32**	.40**	.36**	-.30**	1	
SWM	-.56**	-.52**	-.65**	-.33**	-.44**	-.40**	.26**	-.45**	1

Note. KAIT: Kaufman Adolescent and Adult Intelligence Test, CANTAB: Cambridge Neuropsychological Test Automated Battery; IED: Intra-Extra Dimensional Set Shift; SOC: Stocking of Cambridge; SWM: Spatial Working Memory. * p <.05 ** p <.01

We estimated a three-factor model to test our main hypothesis that Gf, Gc, and EF are correlated. Figure 1 presents the factor model we estimated. All factor loadings were significant (p<.05) and larger than .4. The model fitted the data [$\chi^2(24)$=35.25, p=.07, RMSEA=.050]. SWM loaded highest on EF, followed by SOC and IED respectively, meaning this indicator contributed most to EF and to the EF-Gf relation.

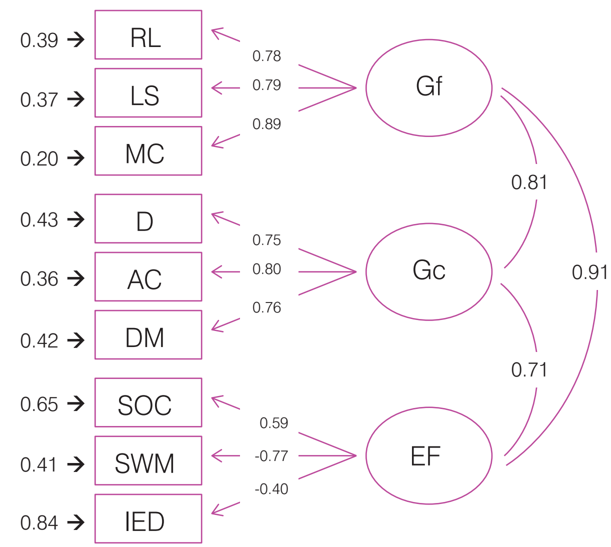


Figure 1 Final model

Note. Structural Equation Modelling examining the relation between Gf and EF. Ellipses represent latent variables; squares represent manifest variables. The curved arrows are correlations between the latent variables. The straight arrows to the right are factor loadings, all significant at the .05 level. The small arrows to the left are residual variances. Residual variances of the latent variables were fixed at 1.00. Negative values are the result of the operationalization of the manifest variables using error scores.
Gf: fluid intelligence, Gc: crystallized intelligence, EF: executive functioning, RL: rebus learning, LS: logical steps, MC: mystery codes, D: definitions, AC: auditory comprehension, DM: double meanings, SOC: stockings of Cambridge, SWM: spatial working memory, IED: intra/extra dimensional set shift.

It is not possible to have a direct test of the hypothesis that the correlation between EF and Gf is higher than the correlation between EF and Gc. Instead, we tested whether the correlation between EF and Gc is equal to the correlation between EF and Gf. If so, then the hypothesis is rejected. If not, then we look at the estimated correlations to draw conclusions. The extra constraints on the model were evaluated with the χ^2 difference test. We used the unrestricted three-factor model, as presented in Figure 1, as our baseline model. The model with equality constraints on the correlations was rejected [$\chi^2(25) = 47.08$, $p = .005$, RMSEA = .069]. The χ^2 difference test [$\chi^2(1) = 11.83$] indicates that the extra constraint results in a significant increase of χ^2.

Given the high correlation between Gf and EF, similar restrictions were applied to test whether they are interchangeable ($r_{Gf\text{-}EF} = 1.00$, $r_{Gc\text{-}Gf} = r_{Gc\text{-}EF}$). Model fit was moderate; [$\chi^2(26) = 39.60$, $p = .04$, RMSEA = .053], indicating that the two are statistically indistinguishable. We further verified the distinctiveness of this strong EF-Gf relation by comparing them to Gc. Constraints were applied to examine equality between Gf and Gc ($r_{Gf\text{-}Gc} = 1.00$, $r_{EF\text{-}Gc} = r_{EF\text{-}Gf}$) and between all three constructs ($r_{Gf} = r_{Gc} = r_{EF}$). Both restricted models did not fit the data: [$\chi^2(26) = 74.79$, $p = .00$, RMSEA = .100] and [$\chi^2(26) = 47.11$, $p = .01$, RMSEA = .066], respectively.

Discussion

The present study examined the relation between EF and Gf in a mixed neuropsychiatric and non-clinical sample using the KAIT as a measure of Gf and Gc, and a selection of CANTAB tasks as a representation of EF. Results showed a significant correlation between Gf and EF, which were statistically indistinguishable in the current model. Working memory was a strong indicator for EF, represented in a high loading of SWM, followed by SOC. Current results are consistent with findings of a strong relation between Gf and working memory (Duncan et al., 2012; Friedman et al., 2006; Redick et al., 2012; Unsworth & Engle, 2006).

Looking at Table 3, SWM shows higher correlations with the KAIT fluid subtests than the other CANTAB tasks do. Although the fluid subscale of the KAIT is assumed to measure a broad scope of cognitive requirements (associative learning, sequencing, planning, syllogistic reasoning, mathematics, hypothetic-deductive reasoning and flexibility), spatial working memory seems to be an essential requirement for an adequate execution of the tasks. Hence, the structure of the KAIT and CANTAB was the starting point of the developed model. An alternative model with the SWM as predictor of Gf was not tested, since it would not contribute to the understanding of either the KAIT or CANTAB. Still, results strengthen the assumption that working memory plays a key role in understanding Gf (Salthouse & Pink, 2008).

Previously, underlying performance of complex cognitive tasks has been referred to as ‘executive attention’ or ‘cognitive control’ (McCabe et al., 2010; Kane, Hambrick, Tuholski, Wilhelm, Payne & Engle, 2004). Comparable, Duncan’s description of the multiple demand (MD) system theory (Duncan et al., 2012, Duncan, 2010) supports the view that EF and Gf share common processes. Essentially, it states that, when performing any set of (complex) actions, a task model is constructed. In this model, task components compete for representation. Adding new components to the model

(e.g., new instructions) leads to more competition, making each component less robust or even lost. The efficiency of constructing such a task model is closely related to Gf, especially when task complexity and novelty increase (Duncan et al., 2012). Since the CANTAB tasks are multi-faceted and increase in complexity compared to singular EF tasks (e.g., go/no-go paradigm), they may require more Gf involvement or MD activity, which in turn may explain the high EF-Gf relation.

Current results have some implications for neuropsychiatric disease and treatment. Clinicians tend to strive for purity of cognitive constructs, which is reflected in commonly used neuropsychological instruments. Not looking at the assemblage of Gf and EF, interaction effects between these cognitive abilities, which are essential in the understanding and explanation of pathological behaviour, will be lost. This is in part due to the fact that most EF tasks are developed from the diverse nature of EF, therefore not focusing on underlying common/general abilities. Following Diamond’s (2013) and Duncan’s theoretical position (Duncan et al., 2012; Duncan, 2010), deficits on task performance do not depend only on separate cognitive task demands, but on their context, i.e., how they are put together to set up goal directed behaviour. Therefore, the assessment of neuropsychological functioning should focus on dissecting the general process and efficiency of rule acquisition and application, next to examining specific cognitive skills necessary for task execution.

Some remarks about task-selection and data collection must be made. First, although CANTAB tasks can be considered a realistic representation of EF in daily life, complexity seems to play such a crucial part in Gf involvement that utilization of less multifaceted EF tasks could have resulted in a different outcome. The amount of general cognitive processes versus specific EF demands required for the tasks will influence this overlap. Second, the inclusion of the IED can be debated given its tendency towards a bimodal distribution. However, mental flexibility is of such importance in defining EF, that exclusion would undermine the a priori formulated model. Third, data collection was based on convenience sampling. Combining subsamples in one group allows us to include both high functioning and impaired participants and to examine the entire scope of EF and intelligence in a heterogeneous sample. A larger dataset may allow future multi-group comparisons, using different psychiatric diagnostic groups and healthy participants, and/or different levels of severity in executive dysfunctioning.

A final comment concerns the theoretical framework. The current study adopts a neuropsychological perspective on EF, and based the model on the Gf-Gc distinction rather than on the more extensive Cattell-Horn-Carroll theory of cognitive abilities (CHC-theory; Schneider & McGrew, 2012). In the latter, purity of abilities is essential in the psychometric perspective on intelligence, whereas the former neuropsychological view tends to be more integrative in describing different interacting abilities. Indeed, according to Kaufman & Kaufman (1993) the fragmentation of intelligence does not contribute to clinical relevance, and therefore, development of the KAIT was only loosely based on CHC theory. For further reading about CHC theory and neuropsychological constructs, see Schneider & McGrew (2012), and Flanagan, Alfonso, Ortiz & Dynda (2013).

In sum, results of the present study strengthen earlier findings on overlap of Gf and EF (Roca et al., 2014, Friedman et al., 2006; Duncan et al., 2008). Following Duncan's theory on the MD system (Duncan, 2010; Duncan et al., 2012), cumulating complexity will lead to more involvement of Gf, and may explain the current high EF-Gf relation. Existing neuropsychological instruments are developed from a multiple-system view and do not separate specific executive task demands and 'higher level' general cognitive control required to execute the task. Furthermore, static outcome measures generally used in neuropsychological assessment do not give insight in the efficiency of task execution. Therefore, Gf dysfunction in neuropsychiatric patients warrants further EF examination and vice versa, to optimally enable discrimination between specific versus general cognitive dysfunctioning. Such a detailed analysis of the process of task execution (using both general intelligence tests as well as neuropsychological instruments) can guide differential diagnosis and lead to a more refined neuropsychiatric treatment indication.

Chapter 5

Predictive Value of Traditional Measures of Executive Functioning on Broad Abilities of the Cattell-Horn-Carroll theory of Cognitive Abilities

Van Aken, L., Van der Heijden, P.T., Oomens, W., Kessels, R.P.C., & Egger, J.I.M. (submitted). Predictive value of traditional measures of executive functioning on broad abilities of the Cattell-Horn-Carroll theory of cognitive abilities.

Abstract

The neuropsychological construct of executive functions (EF), and the psychometric Cattell-Horn-Carroll (CHC) theory of cognitive abilities are both approaches that attempt to describe cognitive functioning. The coherence between EF and CHC abilities is mainly studied using factor-analytical techniques. Through multivariate regression analysis, the current study now assesses the integration of these latent constructs in clinical assessment. The predictive power of six widely used executive tasks on five CHC measures (crystallized and fluid intelligence, visual processing, short-term memory, and processing speed) is examined. Results indicate that executive tasks can predict overall performance on the intelligence tests -except for the Stroop, which only predicts short-term memory and processing speed-, and differentiation in predicting performance between the CHC abilities is limited. It is concluded that executive processes as planning and inhibition are not well represented in intelligence tests. Implications for the use of both EF tests and CHC measures in clinical practice are discussed.

Introduction

In neuropsychology, executive function (EF) refers to multiple higher-order cognitive control processes (Miyake, Friedman, Emerson, Witzki & Howerter, 2000). It is defined as a phenomenon describing the efficiency with which cognitive tasks or demands are handled and problems are solved. No clear consistency is reached as to how many sub functions EF exists of, and literature review reports up to 18 different EF's (Packwood, Hodgetts & Tremblay, 2011). In general, processes like attention, emotion regulation, flexibility, inhibitory control, initiation, organization, planning, self-monitoring, and working memory can be identified as executive processes (Goldstein, Naglieri, Princiotta & Utero, 2014; Naglieri & Goldstein, 2013; Miyake et al., 2000; Packwood et al., 2011). Impairments in these processes may lead to severe dysfunction in a wide range of behaviours (Lezak, Howieson, Bigler & Tranel, 2012). The concept of EF is primarily concerned with information processing and originates from neuropsychological theory such as the central executive hypothesis (Baddeley, 1996; describing a control system with multiple functions) and Luria's (1980) planning, attention, simultaneous, and successive (PASS) information processing model (Das, Naglieri & Kirby, 1994; see also Goldstein et al., 2014).

In the past decades, few of cognitive abilities have been studied more extensively than the Cattell-Horn-Carroll theory of cognitive abilities (CHC theory; McGrew, 2009; Schneider & McGrew, 2012). In CHC theory, the general intelligence factor *g* (Spearman, 1927) stands on top of a hierarchy of cognitive abilities. CHC theory currently identifies sixteen broad abilities each consisting of multiple narrowly defined capacities (Schneider & McGrew, 2012). Of these, six domain-independent general capacities have been described: fluid reasoning (Gf), short-term memory (Gsm), long-term storage and retrieval (Glr), processing speed (Gs), reaction and decision speed (Gt), and psychomotor speed (Gps). In addition, four abilities address acquired knowledge: comprehension-knowledge (Gc), domain-specific knowledge (Gkn), reading and writing (Grw), and quantitative knowledge (Gq). Six further abilities concern domain-specific sensory and motor functions: visual processing (Gv), auditory processing (Ga) and olfactory (Go), tactile (Gh), kinesthetic (Gk) and psychomotor (Gp) abilities.

Both EF and CHC theory are useful in describing cognitive functioning, and both strive for process-pure measurement of cognitive abilities in individuals. In clinical practice, both EF measures and intelligence tests are used simultaneously in neuro-psychological assessments. Both traditions, however, face important issues with respect to operationalization. Especially EF tasks have always been afflicted with task impurity (Miyake et al., 2000). That is, process-pure measurement of EF is hardly possible, as tasks often also rely on other cognitive skills. Consequently, test scores provide no insight into the underlying EF processes. In turn, assessment of intellectual function deals with similar problems. For example, mapping CHC abilities on recently developed intelligence (sub)tests is complex (Benson, Hulac & Kranzler, 2010, Van Aken, Van der Heijden, Hermans, Van der Veld, Kessels & Egger, 2015b; Weiss, Keith, Zhu & Chen, 2013a). Furthermore, it is unclear to what extent subtests are a good representation of the broad abilities. For instance, whereas the working memory index of the WAIS-IV is a valid measure of working memory, it has a too low manipulation load and does not incorporate non-verbal working memory demands, thereby not comprising the full scope of the Gsm factor (Egeland, 2015). Furthermore, factor analysis has shown that most subtests of intelligence batteries are impure measures of CHC abilities, showing cross-loadings with other subtests and latent factors (Grégoire, 2013).

Neuropsychological theory on EF aims to describe cognitive processes that predict and explain (pathological) behaviour, relevant for clinical practice. Therefore, EF has been studied mostly in clinical settings. CHC abilities are mainly derived through factor analysis of large datasets, and have mainly been studied in non-pathological samples. Whereas the neuropsychological approach focuses on predictive and external validity of tests, for instance aiming to predict daily functioning of patients after brain injury, the psychometric approach aims to develop 'pure' measurements of empirically identified latent constructs (internal and structural validity; McGrew, 2010).

The role of neuropsychological constructs, particularly EF, in CHC theory is currently under debate (Jewsbury, Bowden & Duff, 2016). From their respective definitions only, it may be derived that EF and intelligence are partly overlapping constructs. Executive functions operate as 'control mechanisms' for several cognitive processes and have been found to share great variance with general intelligence, facilitating complex behaviour, and (creative) thought (Benedek, Jauk, Sommer, Arendasy & Neubauer, 2014; Floyd, Bergeron, Hamilton, & Parra, 2010; Salthouse, 2005). The relationship between EF and Gf has been studied extensively and is considered strong (Diamond, 2013; Godoy, Dias & Sewabra, 2014; Salthouse, Atkinson & Berish, 2003; Van Aken, Kessels, Wingbermühle, Van der Veld & Egger, 2015a; Duggan & Garcia-Barrera, 2015). Particularly working memory and Gf are strongly related (Chuderski, 2013; Duncan, Schramm, Thompson, & Dumontheil, 2012; Redick, Unsworth, Kelly & Engle, 2012; Salthouse & Pink, 2008; Schneider & McGrew, 2012).

The compatibility of neuropsychology and CHC theory has mainly been studied using factor-analytic modeling. Hoelzle (2008) re-analyzed existing data of seven cognitive domains, including EF, grouping neuropsychological domains into empirically supported CHC domains. In general, the examined attention tests were associated with Gs, executive tests mostly represented Gf and Gv, and most perceptual measures reflected Gv. Memory tests were identified as Glr, and language

tests all included Gc. Floyd et al. (2010) examined the compatibility of EF and CHC theory using the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan & Kramer, 2001) and concluded that subtests of the D-KEFS were just as much a measure of EF as of *g* and the broad CHC abilities. Outcomes of a large study on tests of EF and cognitive abilities and their unique contributions to aging (using both structural modeling and regression techniques) suggested similar conclusions (Salthouse, 2005). The hypothesized EF tasks were found to explain unique variance over the variance explained by the other cognitive abilities. Most EF tasks were closely related to reasoning and perceptual speed tasks, the latter having better psychometric properties than the EF tasks.
Recently, Jewsbury, Bowden & Duff (2016) concluded that the concept of EF is redundant in the CHC model. A separate EF factor in the model did not add to the explained variance by the CHC model. Instead, EF variables could be explained mainly by the Gs, Gv, Gsm, and Gf abilities. In contrast, Roberds (2015) emphasizes the distinctiveness between executive processes and the CHC model. Using canonical correlation and multiple regressions, she demonstrated a large shared amount of variance, but also showed measures of both CHC abilities and EF to have unique variance. Again, Gsm, Gv, Gs, and Gf were identified as the most important CHC abilities in describing the relationship with tests of executive functioning. Recently, Miyake's widely-used switching, inhibition, and updating classification of executive functions (Miyake et al. 2000) was studied in terms of CHC theory. Using confirmatory factor analysis, Jewsbury, Bowden and Strauss et al. (2015) found that switching could be classified as a narrow factor under Gs, and inhibition was entirely explained by Gs, and updating best fitted on the Gsm factor.

Factor-analysis in itself is helpful to unravel the coherence between EF processes and CHC constructs, both by exploring their relation at a latent level, and by studying which observed variables are good indicators of the latent construct. Previously, we also studied latent constructs of EF and intelligence tests through factor analysis, finding overlap mainly between EF and Gf / Gv (Van Aken et al., 2015a; Van Aken et al., 2015b; Van Aken, Kessels, Wingbermühle, Wiltink, Van der Heijden & Egger, 2014). As Schneider & McGrew (2012) state, other (multivariate) statistical techniques should be used next to factor analytic investigations of latent constructs, 'to allow for the simultaneous examination of content (facets), processes, and processing complexity of CHC measures' (p.109). At present, we want to gain more insight in the translation and applicability of these latent constructs to the daily use of tests in clinical practice, adding to the abundant amount of factor-analytical studies on the topic. To do so, the current study adopts a directive approach to examine the possible cohesion and overlap between the operationalizations of EF and CHC theory. It will examine the predictive value of widely-used EF tests for the performance on intelligence tests in a clinical population. Since EF and CHC share similar latent constructs, it can be hypothesized that EF assessment is capable of predicting the level of intellectual functioning (as a reflection of *g*) as well as the outcome on different IQ/index scales as the operationalization of separate CHC abilities. Although one could argue that the direction of this hypothesis can also be reversed (i.e., intelligence as a predictor of EF), the hypothesized direction is chosen since executive processes are more directly related to neurocognitive models, brain functioning and neural efficiency than the psychometric construct of intelligence. Furthermore, using CHC abilities as starting point in explaining variance addresses the critique of overfactoring (Canivez & Kush, 2013; Frazier & Youngstrom, 2007). This critique states that adding complexity to a latent model will lead to better model fit, just because there is more to account for in a complex model, and not because of content arguments .

A set of six traditional and frequently used EF tests will be studied, tapping a wide range of EF abilities that are measured by both singular and more complex executive tasks (Goldstein et al., 2014; Lezak et al., 2012). Hypothesized effects (see Table 1) are based on the previously addressed studies on EF tests and CHC (Hoelzle, 2008; Floyd et al., 2010; Roberds et al., 2015; Salthouse, 2005). Some studies reported conflicting relations, but given the exploratory nature of the current study, we used all positive correlations described by these studies to substantiate our hypotheses. The hypotheses are based on described associations between constructs using the same or similar versions of the tasks.

Table 1 Hypothesized effects of executive tasks on cognitive abilities

	Gc	Gv	Gf	Gsm	Gs
Fluency	+[2]	+[2]	+[2]		+[5]
WCST	+[3]	+[1]	+[1]		
RCFT		+[1]	+[2]		
ToL	+[5]		+[4]	+[4]	
Stroop	+[2]	+[4]	+[1]	+[4]	+[1,3,5]
TMT				+[3]	+[1,3,5]

Note. [1] Hoelzle, 2008, [2] Godoy, 2014, [3] Floyd et al., 2010, [4] Roberds, 2015. [5] Salthouse, 2005. Gc: crystallized intelligence, Gv: visual processing, Gf: fluid intelligence, Gsm: working memory, Gs: processing speed. WCST: Wisconsin Card Sorting Test, RCFT: Rey's Complex Figure Test, ToL: Tower of London, TMT: Trail Making Test.

According to Schneider and McGrew (2012), "CHC based neuropsychological assessment holds great potential. Much clinical lore within the field of neuropsychological assessment is tied to specific tests from specific batteries. CHC theory has

the potential to help neuropsychologists generalize their interpretations beyond specific test batteries and give them greater theoretical unity" (p. 109). The authors argue that the integration of CHC and executive functioning using predictive techniques is relevant for both theorization as well as clinical practice, by bridging the gap between both fields and understanding cognitive processes underlying the CHC abilities.

Method

Participants

Participants were 1,185 in- and outpatients (M_{age} = 35.8 ± 15.0, 57.2% men, M_{IQ} = 88.1 ± 16.9) of the Vincent van Gogh Institute for Psychiatry in Venray, the Netherlands. In accordance with the guidelines of the institutional review board, concealed patient records were drawn from a large electronic database, containing test results of patients registered in the period from January 2004 to November 2015. Data collection was part of a more extensive (neuro)psychological assessment, which generally took place in two sessions of approximately three hours testing time. The majority of patients received psychopharmalogical treatment to relieve symptoms of mental illness. Psychological assessment is not standard procedure during admission to the institute, so only complex comorbid cases with multiple DSM-IV-TR diagnoses (including affective and anxiety disorders, psychotic disorders, impulsivity related disorders, developmental disorders, and comorbidity with personality disorders) were referred for assessment, mostly being psychiatric patients with cognitive complaints and/or impaired cognitive abilities.

Materials

Executive functioning

Included were the Dutch-language versions of the Controlled Oral Word Association Test (category fluency: animal and profession naming [Van der Elst, Van Boxtel, Van Breukelen & Jolles, 2006], n = 620), the Stroop Color Word Test (Stroop; included for analyses: time on card 3 divided by time on card 1 [Stroop, 1935; Van der Elst, Van Boxtel, Van Breukelen & Jolles, 2006], n = 915), the Trail Making Test (TMT; included for analyses: time used for letter-number version -TMT B- divided by time used for number version -TMT A- [Bowie & Harvey, 2006], n = 712), the Wisconsin Card Sorting Test (WCST; included for analyses: total number of cards used [Heaton, Chelune, Talley, Kay & Curtiss, 1993], n = 789), the Tower of London (ToL; included for analyses: total score [Shallice, 1982], n = 747), and Rey's Complex Figure Test (RCFT: included for analyses: total copy score [Rey, 1941; Van der Elst, Van Boxtel, Van Breukelen & Jolles, 2005], n = 833).

Intelligence

All participants completed one out of three intelligence tests; either the Kaufman Adolescent and Adult Intelligence Test (KAIT: Kaufman & Kaufman, 1993; Mulder, Dekker, & Dekker, 2005), or the Dutch-language versions of the third or fourth edition of the Wechsler Adult Intelligence Scale (WAIS-III: Wechsler, 1997, 2000; WAIS-IV: Wechsler, 2008, 2012a, 2012b). The index scores of these intelligence tests are used as operationalizations of CHC constructs. For the KAIT and the WAIS-III, the used IQ/index scores are not a direct translation of the referring CHC construct(s). From now on, therefore, when mentioning the IQ/index score, the corresponding CHC ability will always be pointed out

KAIT (n = 212)

The KAIT is an intelligence test based on, among others, the Gf-Gc theory of Horn & Cattell (1966). A total of six subtests make up two scales; a Crystallized IQ scale (CIQ) and Fluid IQ (FIQ) scale, which are included in the current study as measures of respectively Gc and Gf. Subtest reliability of the Dutch language version of the KAIT (Cronbach's α) varies between .78 and .91 (Mulder et al., 2005).

WAIS-III (n = 779)

Scores of the four indices representing different CHC constructs (Alfonso, Flanagan & Radwan, 2005) were included for analysis: The Verbal Comprehension Index (VCI) representing Gc; the Perceptual Organization Index (POI) as a combined measure of Gv and Gf; the Working Memory Index (WMI) representing Gsm; and the Processing Speed Index (PSI) as a measure of Gs. Split-half reliabilities of the Dutch-language versions of the WAIS-III subtests are similar to the US version and vary between .72 and .93.

WAIS-IV (n = 194)

As for the WAIS-IV, the original index scores (VCI, PRI, WMI, and PSI) were translated to CHC scores (Gc, Gv, Gf, Gsm, and Gs), according to the CHC factors found in previous research (Benson et al., 2010; Van Aken et al., 2015b, Weiss et al., 2013). The CHC scores are a summation of the scaled subtest scores; Gc consists of the four VCI subtests Similarities, Vocabulary, Information, and Comprehension. Gv is made up of the three PRI subtests Block Design, Visual Puzzles, and Picture Completion. Gf consists of the PRI and WMI subtests Matrix Reasoning, Figure Weights, and Arithmetic. Gsm consists of the WMI subtests Digit Span and Letter-Number Sequencing, and Gs is made up of PSI subtests Symbol Search, Coding, and Cancellation. In the case of cross-loadings, the subtest was attributed to the factor of its highest loading (for instance, Arithmetic is part of the WMI and does measure Gsm, but is mainly a test of Gf).

To correctly transform the five CHC (summed scaled subtests) scores into comparable IQ scores, ratio scores were calculated. To do so, the summed CHC scores were corrected for the number of subtests. For instance, the four Gc subtests all contributed 75% to the total factor score, to get a ratio score comparable to the VCI, which was based on three subtests. Next, the ratio scores of Gc, Gsm, and Gs were compared to respectively the VCI, WMI, and PSI converting tables of the WAIS-IV manual. Gf and Gv were compared to the PRI table. Reliability statistics of the Dutch language versions of the WAIS-IV subtests vary between .75 and .93.

Procedures and analysis

Multiple datasets (KAIT, WAIS-III, and WAIS-IV) were used to prevent one-sided interpretation of the CHC constructs. WAIS-III data were included because of collateral data for planning (i.e., the ToL) that were not available in the WAIS-IV data set. Finally, the WAIS-III dataset could more firmly underline conclusions drawn out of results, given the power of the large dataset available.

All tests were administered and scored according to the published test manuals. Each participant completed all executive tasks and one of the three intelligence tests. Data were analyzed using IBM SPSS Statistics 20 (IBM Corp, 2013) and LISREL 8.80 (Jöreskog & Sörbom, 2008). First, the datasets were combined, and a multiple regression analysis was calculated using IBM SPSS Statistics 20 to see whether EF assessment could predict overall level of intellectual functioning. To do so, a total IQ variable (Grand IQ) was made. This dependent variable included all full-scale IQ (FSIQ) scores of the KAIT, WAIS-III, and WAIS-IV to examine global level functioning. Second, multivariate regression analyses were conducted in LISREL 8.8 using the full information maximum likelihood (FIML) estimator. Since each intelligence tests represents similar CHC constructs (each intelligence test has an index representing Gc, for instance), these analyses were executed on the three separate datasets. Standardized estimates (beta-weights) were used to examine the effects of independent variables (executive tests) on the criteria (IQ/index scales). The predictive value of the six executive tests on the CHC abilities was evaluated through three separate multivariate regression analyses using respectively two (CIQ, FIQ; KAIT), four (VBI, POI, WMI, PSI; WAIS-III), and five (Gc, Gv, Gf, Gsm, and Gs; WAIS-IV) criteria per analysis. Due to a large number of missing scores in the WAIS-IV sample on the ToL, this predictor was eliminated for analysis with the WAIS-IV. Other missing values were dealt with according to standard procedures in LISREL 8.8.

Results

Table 2 and 3 show the intercorrelations of all measures. Negative correlations with the WCST are the result the used measure (total number of cards used; higher scores implicate worse performance). First, a multivariate regression analysis was conducted to predict total IQ score (Grand IQ) based on the selection of executive instruments. See Table 4 for details of the model statistics.

Table 2 Correlations between executive tests and intelligence batteries

	Grand_IQ	*KAIT (n=212)*		*WAIS-III (n=779)*				*WAIS-IV (n=194)*				
		CIQ (Gc)	FIQ (Gf)	VCI (Gc)	POI (Gc/Gf)	WMI (Gsm)	PSI (Gs)	Gc	Gv	Gf	Gsm	Gs
Fluency	.53	.50	.56	.52	.37	.46	.51	.49	.45	.51	.53	.67
WCST	-.41	-.34	-.49	-.39	-.45	-.38	-.35	-.50	-.49	-.58	-.36	-.31
RCFT	.53	.36	.49	.50	.57	.50	.46	.48	.42	.53	.45	.15
ToL	.27	.15	.22	.25	.35	.35	.32					
Stroop	.11	.27	.34	.18	.22	.24	.24	.02	.13	.05	.09	.02
TMT	.32	.18	.31	.42	.35	.39	.30	.03	.04	.10	.30	-.04

Note. KAIT: Kaufman Adolescent and Adult Intelligence Test, CIQ: crystallized IQ, FIQ: fluid IQ, WAIS-III: Wechsler Adult Intelligence Scale – third edition, VCI: verbal comprehension index, POI: perceptual organization index, WMI: working memory index, PSI: processing speed index, WAIS-IV: Wechsler Adult Intelligence Scale – fourth edition, Gc: crystallized intelligence, Gv: visual processing, Gf: fluid intelligence, Gsm: working memory, Gs: processing speed. WCST: Wisconsin Card Sorting Test, RCFT: Rey's Complex Figure Test, ToL: Tower of London, TMT: Trail Making Test.

As can be seen in Table 4, all EF tasks except the Stroop significantly contributed to the prediction of overall level of intellectual functioning, with Fluency and RCFT exemplifying strongest effects. Second, three multivariate regression analyses were conducted on each intelligence test. See Table 5 for the standardized solution.

KAIT
Hypotheses of the predictive value of the Fluency, RCFT and WCST were confirmed. The Stroop and ToL had no influence on performance on the CIQ and FIQ, which was contrary to our expectations. Furthermore, the RCFT had an unpredicted significant effect on Gc-performance (CIQ), the TMT did predict Gf-performance (FIQ).

Table 3 Intercorrelations between all independent variables

(n=1185)	Fluency	WCST	RCFT	ToL	Stroop	TMT
Fluency	1.00					
WCST	-.19	1.00				
RCFT	.26	-.39	1.00			
ToL	.14	-.24	.29	1.00		
Stroop	.07	-.10	.10	.23	1.00	
TMT	.18	-.23	.22	.15	.06	1.00
WAIS-IV (n=194)	Gc	Gv	Gf	Gsm	Gs	
Gc	1.00					
Gv	.58	1.00				
Gf	.78	.72	1.00			
Gsm	.63	.47	.65	1.00		
Gs	.56	.63	.60	.59	1.00	
WAIS-III (n=779)	VCI (Gc)	POI (Gf/Gv)	WMI (Gsm)	PSI (Gs)		
VCI (Gc)	1.00					
POI (Gf/Gv)	.68	1.00				
WMI (Gsm)	.76	.66	1.00			
PSI (Gs)	.62	.67	.69	1.00		
KAIT (n=212)	CIQ (Gc)	FIQ (Gf)				
CIQ (Gc)	1.00					
FIQ (Gf)	.74	1.00				

Note. WCST: Wisconsin Card Sorting Test, RCFT: Rey's Complex Figure Test, ToL: Tower of London, TMT: Trail Making Test, WAIS-IV: Wechsler Adult Intelligence Scale – fourth edition, Gc: crystallized intelligence, Gv: visual processing, Gf: fluid intelligence, Gsm: working memory, Gs: processing speed, WAIS-III: Wechsler Adult Intelligence Scale – third edition, VCI: verbal comprehension index, POI: perceptual organization index, WMI: working memory index, PSI: processing speed index, KAIT: Kaufman Adolescent and Adult Intelligence Test, CIQ: crystallized IQ, FIQ: fluid IQ

WAIS-III
All hypothesized effects were confirmed, except for the Stroop, which did not predict performance on the POI, and the ToL, which did not predict VCI performance. Moreover, other effects than those hypothesized were found as well: Fluency predicted outcome on the WMI, and the WCST on the WMI and PSI. The RCFT demonstrated an effect on all indices, while only the POI was hypothesized. A small significant effect of the

Table 4 Linear model with neuropsychological instruments as predictor of the Grand IQ score *(n=1185)*

	Fluency	WCST	RCFT	ToL	Stroop	TMT
b	.58	-.13	1.09	.23	0.71	10.35
SE b	.04	.02	.10	.11	1.92	2.17
B	.38*	-.17*	.31*	.06**	.01	.14*

$R^2 = .50$. * $p < .01$, ** $p < .05$.
Note. b: unstandardized effect, B: standardized effect, WCST: Wisconsin Card Sorting Test, RCFT: Rey's Complex Figure Test, ToL: Tower of London, TMT: Trail Making Test.

Table 5 Standardized total effects of six predictors on performance on the KAIT, WAIS-III, and WAIS-IV IQ/index scales

	KAIT (n = 212)		*WAIS-III (n = 779)*				*WAIS-IV (n = 183)*				
	CIQ (Gc)	FIQ (Gf)	VCI (Gc)	POI (Gf/Gv)	WMI (Gsm)	PSI (Gs)	Gc	Gv	Gf	Gsm	Gs
Fluency	.38*	.37*	.38*	.20*	.31*	.38*	.32*	.29*	.34*	.42*	.66*
WCST	-.17**	-.26*	-.16*	-.21*	-.14*	-.14*	-.28*	-.31*	-.30*	-.06	-.16
RCFT	.18*	.27*	.28*	.35*	.26*	.23*	.25**	.16	.28*	.25*	-.12
ToL	.01	.03	.02	.12*	.14*	.12*					
Stroop	.05	.07	.02	.06	.08**	.09**	-.04	.08	-.03	.05	.01
TMT	.00	.08	.23*	.14*	.19*	.10**	-.06	-.05	-.02	.23*	-.07
R^2	.32	.52	.50	.47	.47	.44	.41	.36	.49	.43	.48

* $p < .01$, ** $p < .05$
Note. KAIT: Kaufman Adolescent and Adult Intelligence Test, CIQ: crystallized IQ, FIQ: fluid IQ, WAIS-III: Wechsler Adult Intelligence Scale – third edition, VCI: verbal comprehension index, POI: perceptual organization index, WMI: working memory index, PSI: processing speed index, WAIS-IV: Wechsler Adult Intelligence Scale – fourth edition, Gc: crystallized intelligence, Gv: visual processing, Gf: fluid intelligence, Gsm: working memory, Gs: processing speed. WCST: Wisconsin Card Sorting Test, RCFT: Rey's Complex Figure Test, ToL: Tower of London, TMT: Trail Making Test.

ToL on PSI next to POI and WMI was found (as was hypothesized), and the TMT also predicted VCI and POI performance besides the predicted effects on WMI and PSI. Predictors with the largest effects were Fluency and RCFT, but they did not differentiate between criteria as hypothesized.

WAIS-IV
Contrary to our expectations, the Stroop did not predict the outcome on any of the CHC factors. Not all hypothesized effects were confirmed. Comparable to the WAIS-III, Fluency was found to be the most important indicator for all CHC factors, although hypothesized to not have an effect on Gsm. All WCST hypotheses were confirmed (i.e., effects on Gf, Gc, and Gv). The TMT predicted Gsm performance, but did not influence Gs outcome as predicted. Lastly, the hypothesized influence on Gv and Gf of the RCFT was not confirmed. The RCFT was related to Gc and Gsm performance, which was unexpected.

Discussion

The goal of the present study was to examine whether EF tasks were able to predict level of intelligence in terms of CHC abilities in a clinical population. This approach is used next to a multitude of latent variable analyses on the topic, in order to investigate the applicability and operationalization of CHC abilities in clinical neuropsychological assessment. Results showed that EF tasks account for 50% of the variance of the overall performance on the IQ tests (see Table 6 for an overview of confirmed and rejected hypotheses, as well as non-hypothesized effects). Except for the Stroop, which did not predict performance on Gf, Gc, and Gv, all hypothesized effects were confirmed as described in Table 1, indicating that EF indeed predicts the general level of intellectual functioning, but that these effects differentiate insufficiently between the CHC abilities. In general, Fluency, RCFT, WCST, and TMT show a similar pattern of effects on all IQ scales, albeit with different magnitudes.

Both Fluency, and RCFT had significant and predictive values for all CHC abilities, explaining on average 37%, and 24% of the variance, respectively. Verbal fluency is known to involve multiple EF processes (Godoy et al. 2014). The same holds for the the RCFT (Van der Meer & Eling, 2008). It can be argued that '*g* saturation' of both Fluency and RCFT is probably high, explaining effects on nearly all CHC constructs. However, the opposite may also be true, looking at the small inter-correlations (except for the correlations of the RCFT with both the Stroop and WCST) between EF tasks and high correlations between IQ scales: the latter share more variance (up to 60%), and mostly appear to be measures of *g,* explaining why broad EF tasks like the Fluency and RCFT do not differentiate between CHC constructs.

Both measures of cognitive flexibility – the WCST and the TMT – also predict performance on all CHC abilities, with 21% and 12% shared variance on average, respectively. However, both tasks are known to measure a broader range of cognitive functions (Nyhus & Barceló, 2009; Sánchez-Cubillo et al., 2009), which may explain the predictive effects of these tests on CHC constructs in the current analysis. The

Table 6 Confirmed, rejected and non-hypothesized effects

	Gc	Gv	Gf	Gsm	Gs
Fluency	+	+	+	n	+
WCST	+	+	+	n	n
RCFT	n	+	+	n	n
ToL	-	n	+	+	n
Stroop	-	-	-	+	+
TMT	n	n	n	+	+

Note. + = confirmed hypothesized effect, - = rejected hypothesized effect, n = non-hypothesized effect. Gc: crystallized intelligence, Gv: visual processing, Gf: fluid intelligence, Gsm: working memory, Gs: processing speed. WCST: Wisconsin Card Sorting Test, RCFT: Rey's Complex Figure Test, ToL: Tower of London, TMT: Trail Making Test.

ToL had small effects on all CHC constructs, except on Gc. Although the ToL is known as a complex planning task covering multiple processes, the current intelligence tests do not seem to adequately capture planning or problem solving skills. The Stroop test is a relatively pure measure of inhibition compared to the other EF tasks, which is considered a core aspect of EF (Miyake et al., 2000). Current results indicate that the Stroop test is capable to specifically predict performance on Gs and Gsm tasks. However, it did not contribute to the prediction of overall intellectual functioning. In addition to all other tasks, the effect of the Stroop on intellectual performance may be negligible, and results especially suggest that any measure of inhibition is absent in the intelligence tests.

The current study has some limitations. The WAIS-III and WAIS-IV differed in their conceptualisation of the indices. The authors suggest the WAIS-IV to be leading in interpreting CHC abilities, since CHC theory motivated development of the WAIS-IV more than the previous version. In doing so, clinicians still should be aware that all IQ tests only tap a small selection of CHC abilities. Furthermore, differences in effects and correlations (see table 2 and 5) between EF and the index scores/IQ scales capturing similar CHC abilities imply diffuse operationalization of theoretical constructs. These discrepancies may be due to variations across subtests and measurement error, resulting in different estimations of the same higher-order factors. In other words, even though similar theoretical constructs are measured, there are considerable differences in their operationalizations. This exposes operationalization difficulties of CHC theory, and emphasizes the need to better capture these abilities. Attempts are being made to further operationalize and purify CHC factors in intelligence tests. For instance, the WISC-V already includes a total of six CHC factors and Gf and Gv are now separate indices (Wechsler, 2014).

The somewhat arbitrary built-up of test indices influences both intercorrelations and extratest correlations of the intelligence tests. The same task-impurity problem holds for the EF tasks, resulting in high measurement error and vague conceptualizations (Miyake et al., 2000; Packwood et al., 2011; Salthouse, 2005). In other words, the concept of EF offers a valid framework in explaining behaviour, but EF test outcomes do not. For instance, a ToL total performance score does not give any information about how the participant planned and organized his behaviour, or if he used adequate strategies to execute the task. Such static outcome measures complicate operationalization of EF, and this plead for more process-oriented assessment techniques.

Another point of discussion is the direction of the analysis. Coming form a neuropsychological perspective, we aimed to examine the capability of EF assessment to predict or identify well-defined (and much researched) CHC constructs, but one could also argue that this should be studied the other way around. Nevertheless, looking at the high correlations between the CHC measures (IQ/index scales) compared to intercorrelations between EF tasks, the current direction seems preferable. Furthermore, EF is more closely related to cognitive models on brain functioning than the empiricially derived CHC abilities.

Current results relate to previous factor analytic research. Clearly, EF tests are most certainly associated with all included intellectual abilities defined by CHC. Floyd et al. (2015) found similar results and concluded that EF tasks are 'too contaminated' by *g*. However, visual inspection of the correlation matrix shows that EF tasks overlap less with each other than the IQ scales, which favors the reverse direction of argumentation. The CHC measures include too much *g,* complicating the differentiation between them, while low correlations between EF tasks reinforce their distinctiveness.

Some studies state that interpretation of IQ tests should be limited to the level of *g,* or the full-scale IQ, and not include the level of the first-order CHC factors (Canivez & Watkins, 2010; Canivez & Kush, 2013). This seems valid looking at current results. Understanding cognitive functioning in patients, however, requires a more fine-grained interpretation. This pleads for supplementary neuropsychological assessment during intelligence testing.

According to Hoelzle (2008), it would be preferrable to use CHC theory as a blueprint for test development rather than as a framework for comprehensively classifying existing measures. We strongly support his view, since our present results confirm that both CHC factors in current intelligence tests and EF tasks are limited in their specificity and are drained by common variance explained by *g.* Depending on the diagnostic questions, clinicians must be aware of the '*g* saturation' of a test and focus on the unique contributions of that test, for instance by correcting for overall intellectual functioning in the normative data, as is already done in most EF tasks.

CHC constructs, and their relation to other (cognitive) theories, are mainly assessed in non-patient populations (Jewsbury, 2016). The current study attempted to investigate the robustness of the used latent CHC constructs and their relation to the (clinically based) construct of EF in a large heterogeneous patient sample. In that respect, future research could further examine current results by investigating different diagnostic groups.

To summarize, the current study adopted a directive and hypothesis-driven approach to clarify the relation between EF and CHC theory and (1) found a high shared variance between CHC measures making it hard for EF tasks to distinguish between separate cognitive abilities, (2) showed that core EF processes were not represented in current intelligence tests, and (3) identified that relations between CHC and EF at a latent level cannot be directly translated to a behavioural level using manifest variables, due to operationalization difficulties in current intellectual and executive assessment.

Results plead for further clarification of both theory and operationalization. Instead of merely being an empirical approach of classifying and describing cognitive abilities on a psychometric level, the CHC model could contribute to the understanding of cognitive processing by expansion into an information processing theory with underpinnings in cognitive neuroscience research. A step in the right direction is suggested by Schneider and McGrew (2012) who introduced a model on "how CHC broad abilities might function as parameters of information processing" (p. 135), which may serve as a framework for future research. In turn, neuropsychological theory of EF is superior in describing cognitive processes, but operationalization of concepts has turned out to be quite difficult (see also Packwood et al., 2011), which could question the applicability of the theory. Neuropsychologists must invest in developing new and better tests to better capture EF processes like planning, and problem solving. For instance, dynamic testing tools are preferrable, since they could contribute to capturing underlying processes (Resing, 2016). Finally, researchers should not only rely on factor analysis to increase insight into the examined constructs (Schneider & McGrew, 2012), but also use other techniques such as non-linear dynamic statistics or network analysis (Borsboom & Cramer, 2013) to examine complex behaviour.

5

Chapter 6

Summary and Discussion

The main objective of this thesis was to gain a better understanding of the relation between EF and the CHC model of intelligence within the framework of contemporary neuropsychological assessment. The studies in this thesis adopted a multi-method approach, and used a clinical neuropsychiatric sample consisting of various psychiatric disorders using various techniques. In this final chapter, an overview of the main results and conclusions will be presented. Furthermore, theoretical and clinical considerations will be discussed, taking the studies' strengths and limitations into account. Finally, directions for future research will be given.

Summary and main findings

Chapter 2 addresses the incorporation of EF in intelligence testing, using the most widely used intelligence test of the past decades, the Wechsler Adult Intelligence Test – Third Edition (WAIS-III; Wechsler, 1997). Although WAIS-III scores are divided into a verbal (VIQ), performal (PIQ), and total IQ (TIQ), a four-factor structure of the WAIS-III was considered best to describe the different intellectual abilities (Kaufman & Lichtenberger, 1999). These factors are labeled as the Verbal Comprehension Index (VCI), the Perceptual Organisation Index (POI), the Working Memory Index (WMI), and the Processing Speed Index (PSI). Even though the framework of the WAIS-III is empirical and pragmatic in nature, the four indices represent five broad abilities that are described in the Cattell-Horn-Carroll theory of cognitive abilities; crystallized intelligence (Gc), fluid intelligence (Gf), visual processing (Gv), short-term / working memory (Gsm), and processing speed (Gs) (Alfonso, Flanagan & Radwan, 2005). The CHC theory can be seen as an extended version of the Gf-Gc theory by Horn & Cattell (1966).

The WAIS-III has been criticized, next to the argument of lacking theoretical ground, to disproportionally assess Gc instead of Gf (Blair, 2006, Duncan et al., 1995). In turn, Gf has been related to EF, which is a core element of neuropsychological assessment. To examine the degree in which EF is represented in the WAIS-III, three EF tests (the Behavioural Assessment of the Dysexecutive Syndrome; BADS, Wilson, Alderman, Burgess, Emslie & Evans, 1996; Krabbendam & Kalff, 1997; the Stroop Color Word Test, Stroop, Hammes, 1971; and the Wisconsin Card Sorting Test; WCST, Heaton, 1981) were included in a WAIS-III factor analysis. Exploratory factor analysis using maximum likelihood procedures was conducted to examine the model fit of two-, three- and four-factor models. The three- and four-factor model fitted the data. The four-factor model fitted the data best, and was almost completely identical to the four-factor indices of the WAIS-III. All three EF tests loaded on factor two, corresponding to the POI of the WAIS-III. Results are in line with the recent findings demonstrating a considerable overlap between EF and Gf, since some subtests of the PIQ scale (which includes the subtests that represent POI) are associated with Gf (Duncan, 2010; Roca et al., 2010). Further investigation of the relationship between EF and Gf will be discussed later on in this chapter. In addition, our results confirm the overload of Gc in the WAIS-III subtests, with VCI accounting for 41.7% of the variance in the model, as opposed to 12.9% of the POI, 6.2% of the WMI, and 5.2% of the PSI. This may result in a biased perspective on intelligence (Blair, 2006). It also questions the content validity of the WAIS-III, particularly considering the fact that in most datasets, *g* is better explained by tests that cover fluid abilities than by tests that cover crystallized abilities (i.e., Gc; Duncan, 1995, 2010).

The latest version of the Wechsler intelligence scale, the WAIS-IV, was examined in **chapter 3**. In the development of the fourth edition of the Wechsler scales, the WAIS-IV, CHC theory guided adjustments of its subtests, especially aimed at improving the measurement of Gf. The four indices have been preserved for the most part, except for POI which is labeled the Perceptual Reasoning Index in the WAIS-IV, to better cover the fluid processes (Gf) assumed to be captured by this index. The subtest Digit Span of the WMI was improved by adding a sequencing condition, to better represent the CHC factor Gsm. Cancellation was added as an optional measure of the PSI, covering Gs processes. The VCI remained relatively unchanged (except for some item changes), and is considered a measure of Gc (Wechsler, 2008). The five CHC abilities previously identified in the WAIS-III now formed the theoretical framework underlying the development and adjustments of the subtests of the WAIS-IV (Benson et al., 2010; Weiss et al., 2013a).

Using confirmatory factor analyses, a WAIS-IV five factor structure (according to the five incorporated CHC constructs) was examined, as an alternative for the four factor structure of the WAIS-IV manual. Main critique on the four-factor structure is that the PRI does not follow CHC structure. Instead, we divided the PRI into two separate CHC factors reflecting Gv and Gf. Gv consists of Block Design, Matrix Reasoning, Visual Puzzles, Figure Weights, and Picture Completion, and Gf consists of the subtests Matrix Reasoning, Figure Weights, and Arithmetic. Although the four-factor structure of the manual fitted the data, the CHC based five-factor model also showed an adequate fit. Despite the principle of parsimony, this CHC-based five-factor model is preferred over the more economic four-factor model, given its theoretical underpinnings. Besides the PRI, all other indices corresponded directly to the other three CHC abilities represented in the WAIS-IV; VCI, WMI and PSI represent Gc, Gsm, and Gs, respectively. Remarkably, the share of Gf in the WAIS-IV again was found to be low. Gf was covered by three subtests in total, only two of which are part of the core battery. In conclusion, the adjustment of this latest version does not fully live up to the statement (and additional expectations) that the measurement of fluid

reasoning is improved. Nevertheless, the five-factor model including Gf is valid, and separation of the Gv and Gf factors enhances understanding of performance on the corresponding subtests.

Zooming in to CHC theory, the Gc-Gf dichotomy is subject of **chapter 4**. This dichotomy can be seen as the core element of the CHC model, with both constructs closest related to *g* compared to other CHC constructs, and also with each other (Benson et al., 2010; Horn & Cattell, 1966; Grégoire, 2013; Schweizer, Troche & Rammsayer, 2011; Schneider & McGrew, 2012). Taking a neuropsychological perspective, Gf shows great resemblance to EF, and the uniqueness of this relation is examined in this chapter. At first glance, definitions of EF are highly similar to Gf. Both are considered to be involved in the execution of complex and efficient behaviour in new or non-automated situations, and overlap between the two has been demonstrated in previous research. Chapter 2 already suggested results in line with this overlap. However, the representation of Gf in the Wechsler scales is limited. Therefore, the KAIT was selected to measure both Gf and Gc, and a latent-factor model was developed to test interrelations between Gf, Gc, and EF. Three executive tests of the Cambridge Neuropsychological Test Automated Battery (CANTAB) covering working memory, planning, and set shifting were selected as a representation of EF. Results demonstrated a high overlap between Gf and EF (correlation of .91), different from their relation with Gc. Working memory plays a key role in this relation, reflected in a high contribution of the working memory test of the CANTAB to the Gf-EF relation.

These results are consistent with previous findings on the relation between EF and Gf. The fact that the overlap is particularly high in this study may rely on the selected tests. The amount of general cognitive processes versus EF-specific demands required for adequate performance on the selected task may affect the overlap with Gf, which is considered a general, domain-independent ability (Duncan, 2012). EF tests come in many different shapes, and the selected CANTAB tests are broad and complex in nature compared to reaction-time tests or less complex shifting/inhibition tests like a go/no-go paradigm, influencing the relation with Gf. As Diamond (2013) and Duncan (2010, 2012) state, deficits on task performance do not depend only on separate cognitive task demands, but on their context and interplay, that is, how separate task demands are put together and interact with each other to set up goal-directed behaviour. Taking these considerations into account, it is concluded that general and specific processes must be distinguished when assessing EF.

Intelligence testing, using batteries based on the CHC model, is an important part of neuropsychological assessment. In **chapter 5**, the relations between EF as measured in neuropsychological assessment and CHC domains as measured with intelligence tests were examined. The previous studies used factor analysis to examine correspondence between the latent constructs EF and CHC domains intelligence and how they were represented in neuropsychological and intelligence tests. Still, the question remains whether these latent constructs are adequately operationalized. Therefore, a series of widely used executive tasks were examined on their predictive value on the outcome of CHC domains as measured with intelligence tests.

Multivariate regression techniques were applied to examine the predictive value of EF tasks on the CHC domains. Results showed that EF tasks accounted for, on average, 50% of the variance in the performance on the IQ tests. Current results relate to previous factor analytic research. Low correlations between EF tasks in the current study reinforce their distinctiveness, while high shared variance between the CHC measures point to a high ‘g saturation’ of these measures.

Some EF tests, notably planning and inhibition tests, did not predict performance on any of the intelligence scales, suggesting a lack of coherence between inhibition and planning skills and current intelligence measures. Alternatively, this may also be due to the poor psychometric properties and task impurity of (especially) the ToL. Unfortunately, this problem applies to several other EF tests as well (Miyake et al., 2000; Salthouse, 2005). It can be concluded that the identified relations between EF and CHC constructs at a latent level cannot be directly translated to a behavioural level using manifest variables, due to operationalisation difficulties in both executive and intellectual assessment.

To wrap up, the abovementioned results can be summarized in the following points. **Firstly**, results demonstrated a large overlap between EF and intelligence. Especially EF and Gf shared essential common processes. However, the two constructs are not identical, and the overlap found between several tasks of EF and intelligence depended on the complexity of the selected task. **Secondly**, the neuropsychological perspective on EF and the psychometric (CHC-based) approach both have advantages in describing cognitive functioning, and both may benefit from each other. Neuropsychology has solid theories on EF, but has not been able to adequately translate information processing theories into valid and reliable instruments as is done with the CHC model. CHC theory could evolve by integrating neurocognitive processing into the model as is done in neuropsychological theories. **Thirdly**, there is a gap between theories on EF and intelligence and measurement of both constructs in clinical practice. In neuropsychological assessment, clinical interpretation of EF tests must take both general and specific abilities into account, and take notion of the overlap between measures. This prevents a fragmented view on executive abilities in which general abilities are not taken into account. Current intelligence tests based on psychometric models lack a clear relation to theory, and incorporation of executive

processes in intelligence tests is limited. Next, theoretical implications of the strong Gf-EF relation will be discussed, as well as how both EF theories and the CHC model can facilitate each other. In addition, clinical implications will be mentioned.

General discussion

The Gf – EF relation

Results in chapter 3 reflect the limited measurement of Gf in the WAIS-III and WAIS-IV, which is remarkable and not in agreement with the prominent place of Gf in explaining *g* in CHC theory. Furthermore, Gf shows very high overlap with EF, demonstrated in chapter 4. It becomes clear that Gf is the core of what we call intelligent or complex behaviour. Unfortunately, only two subtests of the WAIS-IV core battery seem to capture Gf, both belonging to separate indices of the test (i.e., PRI and WMI), according to the test manual. At the same time, these Gf subtests also show cross loadings with other CHC constructs.

It is clear that Gf is central to the understanding of both the construct of intelligence and the concept of executive function. In the CHC model, Gf is considered a general, domain-independent process (Schneider & McGrew, 2012), involved in new situations where no other learned or automated responses are available or sufficient. This obviously relates to the Multiple Demand system previously described in the general introduction (Duncan, 2010, 2012). Within this system, Gf is seen as the efficiency in which any new or complex task is executed, regardless of other specific cognitive (executive) demands of the task. The Multiple Demand system has a strong relation especially with EF, since executive processes are also involved in tasks which require controlled (as opposed to automatic) responses, but this does not mean that both Gf and EF processes are identical.

Both Schneider & McGrew (2012) and Duncan (2010, 2012) imply that all executive tests contain a general fluid component that is independent of task modality. Unfortunately, this fluid part is hard to identify. Pure measurement of Gf is not feasible without addressing other abilities. Consequently, a Gf task is always ‘contaminated’ with domain-dependent cognitive processes (Grégoire, 2013). If Gf is the general capacity to solve complex problems on top of the cognitive demands of that specific task, it should be measured and identified that way. This implies that in clinical practice there should be more focus on dissecting general versus specific task demands.

The CHC model and executive function

Referring to chapter 3, the CHC model is represented in current intelligence tests, but not perfectly. The fact that the five-factor CHC structure demonstrates the best model fit in our study does not imply that the proposed original four-factor structure of the WAIS-IV is not valid. However, clinicians must be aware that the PRI consists of both Gv and Gf, according to CHC. Furthermore, WAIS-IV indices are not direct and independent translations of CHC constructs, and the coherence and overlap between factors should therefore be taken into consideration by the clinician interpreting WAIS-IV performance.

The CHC model and neuropsychological theory on EF have developed into approaches describing cognitive processes. Neurocognitive processes underlying EF are described in neuropsychological theories based on clinical relevance in pathological behaviour. With origins in academic (test) psychology, CHC on the other hand intends to describe a factor pure model of different cognitive abilities. Ideally, these two principles should be integrated to have the best of both world; that is, making it possible to identify pure defined (latent) factors that can predict clinically relevant (pathological) behaviour. As is seen in chapter 5, however, the compatibility of CHC constructs and EF measurements is blurred due to *g* saturation of the IQ scales and measurement error of the EF tests. In other words, there is no complete coherence with theoretical constructs and the instruments that are assumed to measure those constructs. This 'gap' between theory and practice leads to several difficulties. Unclear EF operationalizations have led some researchers to conclude that the construct of EF is completely redundant with CHC, especially with Gf, Gv, Gs, and Gsm (Jewsbury et al. 2016), even though no theory on information processing is incorporated in CHC. Others criticize that the interpretation of intelligence tests at the level of the indices (for instance, the VCI as conceptualization of Gc) is not valid at all, since the contribution of these indices in addition to the variance explained by *g* (or FSIQ in this matter) in intelligence batteries is low (Canivez & Watkins, 2010; Gignac, 2008; Gignac & Watkins, 2013). However, in clinical practice, one must look beyond *g* in order to understand the performance on an intelligence test in patients with cognitive disorders or psychopathology, for instance to explain disharmonic profiles in patients with traumatic brain injury.

Given its relatively straightforward conceptualization, it is tempting to take a model like CHC as a guiding principle in test development. However, although a psychometric, trait-based approach excels in *identifying* different aspects or dimensions of cognition, it is an insufficient model for *explaining* behaviour. In other words, in order to unravel the cognitive processing underlying neuropsychological tests, a purely psychometric approach is insufficient. From this perspective, neuropsychological theory on EF offers a more comprehensive framework in explaining human behaviour.

Unfortunately, the translation of executive processing into valid measurements is easier said than done, and some neuropsychological theories have never been implemented in the field of neuropsychological assessment. Furthermore, the most

widespread neuropsychological tests used to map cognitive (or executive) deficits originate from the first half of the twentieth century, developed using principles that are currently outdated. These tests principally measure the outcome of a cognitive process, instead of capturing this process itself. The gap between science and clinical practice is preserved by holding on to these traditional tests.

Clinical considerations

Current results may guide the design of the neuropsychological assessment. Clearly, the translation from a theoretical construct to the score on a test is complex. This has implications for clinical assessment of EF and intelligence, and of course for test development.

Looking at chapter 2 and 4, it becomes clear that assessment of EF in intelligence tests is limited. Executive dysfunction, for instance, may contribute to impaired performance on POI of the WAIS-III and, as a result, 'disharmonic' distributions of intellectual capacities measured by the WAIS. Therefore, limited POI scores give direction to supplementary EF assessment to further explain this impaired performance on POI. The same holds for the uneven contribution of abilities represented in the WAIS-IV, in which (especially) Gf measurement is poorly represented, which makes additional neuropsychological assessment necessary. At the same time, neuropsychological assessment does not include integrative tests to measure overall (fluid) intellectual capacities. When developing intelligence tests, the role of Gf should be even more prominent given the important role it plays in complex, intelligent behaviour. At the same time, including the measurement of executive processes in these intelligence tests may increase its clinical utility and aid neuropsychological interpretation.

In adopting theoretical constructs like CHC, clinicians must be aware of the fact that tests are no direct representation of the underlying theoretical constructs. For instance, the WMI of the WAIS-IV is a valid measure of working memory, but it is only a limited representation of the working memory construct (or Gsm in CHC terminology); it has a low manipulation load and lacks visuospatial demands (Egeland, 2015). Furthermore, executive tests representing processes like shifting, updating, or inhibition are too fragmented. All these separate executive tests feed the expectation that all tests reflect unique theoretical constructs. This has resulted in a narrow and one-sided interpretation of test performances, ignoring the overlap between measures and measurement error, but also ignoring the overlap in theoretical constructs that are represented.

As is discussed in chapter 4, the overlap between Gf and EF relies on the complexity of a task and the amount of general and specific cognitive abilities that are represented in a task. The more complex a test, the more Gf influence (i.e., activity of the MD system), and the more overlap with EF. Specific executive test demands can be associated with processes outside the MD system, however, execution of current traditional EF tasks like the Wisconsin Card Sorting Test or Verbal Fluency is mainly based on 'general' MD activity (or to put it in other words, the tests are saturated with *g*) instead of these specific regions outside the system that the tests intend to measure (Duncan, 2012). Neuropsychologists have a task in future test development to develop an integrated measurement in which general and specific cognitive abilities are identified and dissected in one task. This is preferably done using a 'complexity-gradient' that can better distinguish general capacities (explained by Gf impairment) from specific executive test demands. This approach needs process-oriented assessment techniques (see also Resing, 2016) instead of mapping multiple static outcomes, which is currently routine. Seen from the psychometric view, the CHC model could be integrated in a new, process-oriented EF approach, and CHC based neuropsychological assessment could increase external validity, by using a comprehensive framework and using a common nomenclature. A first step coming from CHC was suggested by Schneider and McGrew (2012), who introduced a model on 'how CHC broad abilities might function as parameters of information processing; (p.135), which in turn may serve as a framework for future research. In sum, analysing cognitive processing on a behavioural level in terms of general and specific disabilities can guide differential diagnosis and lead to a more refined neuropsychiatric treatment indication.

Limitations

A possible limitation of the current studies is the use of a heterogeneous neuropsychiatric group obtained by convenience sampling. This may have resulted in lowered average scores and high scatter. Psychopathology, medical treatment, mental effort, motivation or perceived failure may all affect test performance and score profiles. For instance, higher EF intercorrelations and relations with intelligence measures are assumed in clinical populations with (actual or assumed) frontal-lobe dysfunction (Friedman et al., 2006; Rabbit, Lowe & Shilling, 2001). It is unknown whether current results generalize to healthy samples or to other impaired populations (brain injured patients for instance) or age ranges (children, elderly). However, measurement invariance among healthy and clinical samples is demonstrated for the WAIS-III (Van der Heijden & Donders, 2003a), and the WAIS-IV (Weiss et al., 2013a). The factor structures identified in chapters 2, 3, and 4 were comparable to the facture structures in the normative samples, and no group differences in correlational structures (between EF, Gf, and Gc) were found between the healthy controls and the psychiatric

patients in chapter 3. These results, in combination with the heterogeneity of the sample, confirm the robustness of the factor structures used in the current studies. Furthermore, the application of CHC theory in clinical samples is limited, even though intellectual assessment is often part of neuropsychological assessment. Therefore, it is highly important that the model is examined in the clinical field. First results are promising, demonstrating measurement invariance of the CHC model between healthy and clinical groups (Jewsbury et al., 2016). Future studies should further differentiate these findings, particularly by examining the relation between EF and intelligence within a CHC framework in other patient samples and using multi-group comparisons, enhancing the external validity.

Intelligence theory has mainly been driven by a factor analytic approach, which is also the main method of the present thesis. As discussed before, factor analysis is mainly pragmatic in origin, despite using CHC as a theoretical framework. It is still an open question whether identified factors truly represent underlying dimensions. Furthermore, factor analysis has identified a growing number of factors, and operationalization often leads to 'over-factoring' (Frazier & Youngstrom, 2007). As argued before, this growing enumeration of factors does not per se contribute to the neuropsychologists understanding and explanation of (pathological) behaviour (Blair, 2010).

The clinical field will not move away from pure classification and enumeration of cognitive abilities if factor analysis remains the main technique. The discussion on the pros and cons of factor analysis is larger than the scope of this thesis, but aligns with the conclusions that clinicians need to focus on capturing general and specific cognitive processes in neuropsychological assessment. In addition to factor analysis, other regression techniques (as is done in chapter 5) or alternative techniques such as network analysis (see also Borsboom & Cramer, 2013) can be used to enhance understanding of these processes.

Conclusion and future directions

In clinical practice, methods must be based on theoretical foundations. Clinical assessment has driven away from solid theory, resulting in an increasing distance between approaches, an abundancy of hypothetical constructs and neuropsychological tests, as well as task-driven clinical interpretation. As a result, we no longer can see the wood for the trees. A core task of psychologists is being aware that no single theory captures the characteristics of every individual or every test, but this is by no means reason to abandon theory altogether. In contrast, this awareness should stimulate and motivate the clinician and the scientist to handle operational shortcomings and to understand individuals within the framework of available theories.

Both the CHC model and neuropsychology have been valuable for the understanding of human cognition. In general, both seem to reach a similar level of explanation with regard to human behaviour. In clinical practice, it is crucial to formulate hypotheses based on theoretical constructs that can be measured. Psychometric techniques are superior in identifying coherence in cognitive tests at a latent level and translating them into valid measurements, but latent constructs do not contribute to the *understanding* of cognitive processes per se. Cognitive models in neuropsychology, on the other hand, may be too strict in describing cognitive processes, leading to difficulties in operationalization of the theoretical constructs. This limits the usefulness of both approaches, and solutions are not easily handed.

The gap between executive processes and CHC abilities will probably persist when examined using traditional approaches, datasets and techniques. In other words, psychologists are used to working with linear techniques, while it is evident that behaviour, especially on the extremes of the bell curve, is a non-linear phenomenon. Linear models are necessary to create understanding of constructs and their coherence at group level, but they only to some extent contribute to the assessment of impaired individuals. Instead, non-linear statistics and network analysis might be helpful to assess information processing models in individuals or interaction between constructs, respectively. These approaches provide techniques other than a latent variable approach to explain complex behaviour, thereby providing a promising alternative psychometric lens on cognitive functioning.

References

References

A

American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders,* Fourth Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000.

Anderson, T.W., & Amemiya, Y. (1988). The asymptotic normal distribution of estimators in factor analysis under general condition. *The Annals of Statistics, 16,* 759-771. doi: 10.1214/aos/1176350834

Alfonso, V. C., Flanagan, D. P., & Radwan, S. (2005). The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In D.P. Flanagan & P.L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 185-202). New York, NY, US: Guilford Press.

Ardilla, A. (1999). A neuropsychological approach to intelligence. *Neuropsychology Review, 9,* 117-136.

Arnau, R.C., & Thompson, B. (2000). Second-order confirmatory factor analysis of the WAIS–III. *Assessment, 7,* 237-246. doi: 10.1177/107319110000700304

B

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology: Section A*, *49*, 5-28. doi: 10.1080/713755608

Baddeley, A.D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4,* 829-838.

Baddeley, A.D. (2007). *Working memory, thought and action.* Oxford: Oxford University Press

Barbey, A.K., Colom, R., Paul, E.J., & Grafman, J. (2014). Architecture of fluid intelligence and working memory revealed by lesion mapping. *Brain Structure and Function, 219,* 485-494. doi: 10.1007/s00429-013-0512-z

Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. C. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. *Intelligence, 46*, 73-83. http://dx.doi.org/10.1016/j.intell.2014.05.007

Benson, N., Hulac, D. M., & Kranzler, J. H. (2010). Independent examination of the Wechsler Intelligence Scale – Fourth Edition (WAIS-IV): What does the WAIS-IV measure? *Psychological Assessment, 22,* 121-130. doi: 10.1037/a0017767

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioural and Brain Sciences, 29,* 109-160.

Blair, C. (2010). Fluid cognitive abilities and general intelligence. In W.F. Overton & R.M. Lerner (Eds). *The handbook of life-span development: Volume 1. Cognition, biology, and methods.* (pp. 226-258). New York, NY, US: Wiley.

Borsboom, D., & Cramer, A. O. (2013). Network analysis: an integrative approach to the structure of psychopathology. *Annual review of clinical psychology, 9*, 91-121. doi: 10.1146/annurev-clinpsy-050212-185608

Bowden, S.C. (2013). Theoretical Convergence in Assessment of Cognition. *Journal of Psychoeducational Assessment, 31,* 148-156. doi: 10.1177/0734282913478035

Bowie, C.R., & Harvey, P.D. (2006). Administration and interpretation of the Trail Making Test. *Nature Protocols, 1,* 2277-2281. doi: 10.1038/nprot.2006.390

Byrne, B. M. (2001). Structural equation modeling: Perspectives on the present and the future. *International Journal of Testing, 1,* 327-334. doi: 10.1080/15305058.2001.9669479

C

Canivez, G. L., & Kush, J. C. (2013). WAIS-IV and WISC-IV structural validity: Alternate methods, alternate results commentary on Weiss et al. (2013a) and Weiss et al. (2013b). *Journal of Psychoeducational Assessment, 32,* 1-13. doi: 10.1177/0734282913478036

Canivez, G. L., & Watkins, M. W. (2010). Investigation of the factor structure of the Wechsler Adult Intelligence Scale –Fourth Edition (WAIS-IV): Exploratory and higher order factor analyses. *Psychological Assessment, 22,* 827-836. doi: 10.1037/a0020429

Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Chuderski, A. (2013). When are fluid intelligence and working memory isomorphic and when are they not? *Intelligence, 41,* 244-262. doi: 10.1016/j.intell.2013.04.003

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159. doi: 10.1037/0033- 2909.112.1.155

Cox, S.R., MacPherson, S.E., Ferguson, K.J., Nissan, J., Royle, N.A., MacLullich, A.M.J., Wardlaw, J.M., & Deary, I.J. (2014). Correlational structure of ‘frontal’ tests and intelligence tests indicaties two components with asymmetrical neurostructural correlates in old age. *Intelligence, 46,* 94-106.

D

Das, J.P., Naglieri, J.A., & Kirby, J.R. (1994) *Assessment of cognitive processes: The PASS theory of intelligence.* Needham Heights, MA, US: Allyn & Bacon.

Davis, A.S., & Pierson, E.E. (2012). The Relationship Between the WAIS-III Digit Symbol Coding and Executive Functioning. *Applied Neuropsychology: Adult, 19,* 192-197. doi: 10.1080/09084282.2011.643958

Davis, A.S., Pierson, E.E., & Holmes Finch, W. (2011). A Canonical Correlation Analysis of Intelligence and Executive Functioning. *Applied Neuropsychology, 18,* 61-68. doi: 10.1080/09084282.2010.523392

Dekker, R., Mulder, J., & Dekker, P. (2005) Nederlandstalige bewerking van de Kaufman Adolescent and Adult Intelligence test: KAIT. *De Psycholoog, 40*, 451-457.

Delis, D.C., Kaplan, E. & Kramer, J.H. (2001). *Delis-Kaplan executive function system.* San Antonio, TX: Psychological Corporation.

Diamond, A. (2013). Executive Functions. *Annual Review Psychology, 64,* 135-168. doi: 10.1146/annurev-psych-113011-143750

Duggan, E.C., & Garcia-Barrera, M.A. (2015). Executive Functioning and Intelligence. In S. Goldstein, D. Princiotta, & J.A. Naglieri (Eds.), *Handbook of Intelligence* (pp. 435- 458). New York: Springer

Duncan, J. (2010). *How intelligence happens.* New Haven: Yale University Press.

Duncan, J. (2013). The structure of cognition: attentional episodes in mind and brain. *Neuron, 2,* 35-50. doi: 10.1016/j.neuron.2013.09.015

Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia, 33*, 261-268. doi: 10.1016/0028-3932(94)00124-8

Duncan, J., Parr, A., Woolgar, A., Thompson, R., Brigth, P., Cox, S., Bishop, S., & Nimmo Smith, I. (2008). Goal neglect and Spearman’s *g*: competing parts of a complex task. *Journal of Experimental Psychology: General, 137,* 131-148. doi: http://dx.doi.org/10.1037/0096-3445.137.1.131.supp

Duncan, J., Schramm, M., Thompson, R., & Dumontheil, I. (2012). Task rules, working memory, and fluid intelligence. *Psychonomic bulletin & review, 19,* 864-870. doi: 10.3758/s13423-012-0225-y

E

Egeland, J. (2015). Measuring working memory with digit span and the letter-number sequencing subtests from the WAIS-IV: too low manipulation load and risk for underestimating modality effects. *Applied Neuropsychology: Adult, 22,* 445-451. doi: 10.1080/23279095.2014.992069

Eling, P. (2016). Hoe meet je nu intelligentie? *Tijdschrift voor Neuropsychologie, 3,* 45-51.

Evers, A., Muniz, J., Bartram, D., Boben, D., Egeland, J., Fernández-Hermida, J.R., … & Urbánek, T. (2012). Testing practices in the 21st century: developments and European psychologists’ opinions. *European Psychologist, 17,* 300-319. doi: 10.1027/1016-9040/a000102

F

Field, A. (2009). *Discovering statistics using SPSS.* Londen: SAGE Publications Ltd.

Fisher, D.C., Ledbetter, M.F., Cohen, N.J., Marmor, D., & Tulsky, D.S. (2000). WAIS-III and WMS-III profiles of mildly to severely brain-injured patients. *Applied Neuropsychology, 7,* 126-132. doi: 10.1207/S15324826AN0703_2

Flanagan, D.P., Alfonso, V.C., Ortiz, S.O., & Dynda, A.M. (2013). Cognitive Assessment: Progress in Psychometric Theories of Intelligence, the Structure of Cognitive Ability Tests, and Interpretive Approaches to Cognitive Test Performance. In: Saklofske, D.H., Reynolds, C.R., & Schwean, V.L. (Eds). *The Oxford Handbook of Child Psychological Assessment* (pp. 239-285). Oxford: OUP USA.

Floyd, R.G., Bergeron, R., Hamilton, G., & Parra, G.R. (2010). How do executive functions fit with the Cattell-Horn-Carroll model? Some evidence from a joint factor analysis of the Delis-Kaplan Executive Function System and the Woodcock-Johnson III Test of Cognitive Abilities. *Psychology in the Schools, 47,* 721-738. doi: 10.1002/pits.20500

Frazier, T.W., & Youngstrom, E.A. (2007). Historical increase in the number of factors measured by commercial tests of cognitive ability: Are we overfactoring? *Intelligence, 35,* 169-182. doi: 10.1016/j.intell.2006.07.002

Friedman, N.P., Miyake, A., Corley, R.P., Young, S.E., DeFries, J.C., & Hewitt, J.K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17,* 172-179. doi: 10.1111/j.1467-9280.2006.01681.x

G

Gardner, M.K. (2012). Theories of Intelligence. In M.A. Bray & T.J. Kehle (Eds.) *Oxford Handbook of School Psychology* (pp. 79-100). New York: Oxford University Press.

Gignac, G.E. (2008). Higher-order models versus direct hierarchical models: *g* as superordinate or breadth factor? *Psychology Science Quarterly, 50,* 21-43.

Gignac, G.E., & Watkins, M.W. (2013). Bifactor Modeling and the Estimation of Model Based Reliability in the WAIS-IV. *Multivariate Behavioral Research, 48,* 639-662. doi: 10.1080/00273171.2013.804398

Godoy, S., Dias, N.M., & Seabra, A.G. (2014). Executive and Non-Executive Cognitive Abilities in Teenagers: Differences as a Function of Intelligence. Psychology, 5, 2018- 2032. doi: 10.4236/psych.2014.518205

Golay, P., Reverte, I., Rossier, J., Favez, N., & Lecerf, T. (2012). Further insights on the French WISC–IV factor structure through Bayesian Structural Equation Modeling. *Psychological Assessment*. doi: 10.1037/a0030676

Goldstein, S., Naglieri, J.A., Princiotta, D., & Otero, T.M. (2014). Introduction: A History of Executive Functioning as a Theoretical and Clinical Construct. In S. Goldstein and J.A. Naglieri (Eds.), *Handbook of Executive Functioning* (pp. 3-12). New York, NY: Springer.

Grégoire, J. (2013). Measuring components of intelligence: mission impossible? *Journal of Psychoeducational Assessment, 31,* 138-147. doi: 10.1177/0734282913478034

H

Hammes, J.G.W. (1971). *De Stroop kleur-woord test, handleiding* [*The Stroop-Color-Word Test, manual*]. Amsterdam: Pearson Assessment and Information B.V

Hawkins, K.A. (1998). Indicators of brain dysfunction derived from graphic representations of the WAIS–III / WMS–III technical manual samples: A preliminary approach to clinical utility. *The Clinical Neuropsychologist, 12,* 535-551. doi: 10.1076/clin.12.4.535.7236

Heaton, R. (1981). *Wisconsin Card Sorting Test: Manual.* Odessa, FL: Psychological Assessment Resources.

Heaton, R.K., Chelune, G.J., Talley, J.L., Kay, G.G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test Manual. Revised and expanded.* Odessa (FL): Psychological Assessment Resources.

Hoelzle, J.B. (2008). *Neuropsychological assessment and the Cattell-Horn-Carroll (CHC) cognitive abilities model* (Theses and Dissertations). Available from The University of Toledo Digital Repository. (UMI No. 3328214)

Horn, J.L. (1989). Cognitive diversity: A framework of learning. In Ackerman, P.L., Sternberg, R.J., Glaser, R. (Eds.), *Learning and individual differences* (pp. 61-116). New York: Freeman.

Horn, J.L., & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology, 57*, 253-270. doi: 10.1037/h0023816

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6,* 1-55. doi: 10.1080/10705519909540118

I

IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

J

Jaeggi, S.M., Buschkuehl, M., Jonides, J., & Perrig, W.J. (2008). Improving fluid intelligence with training on working memory. *PNAS, 105,* 6829-6833. doi: 10.1073/pnas.0801268105

Jewsbury, P.A., Bowden, S.C., & Duff, K. (2016). The Cattell-Horn-Carroll Model of cognition for clinical assessment. *Journal of psychoeducational Assessment,* 1-21. doi: 10.1177/0734282916651360

Jewsbury, P.A., Bowden, S.C., & Strauss, M.E. (2015). Integrating the switching, inhibition, and updating model of executive function with the Cattell-Horn-Carroll model. *Journal of Experimental Psychology: General.* Advance online publication. doi: 10.1037/xge0000119

Jöreskog, K.G., & Sörbom, D. (2008). LISREL 8.80 for Windows [Computer Software]. Lincolnwood, IL: Scientific Software International, Inc.

Jung, R. E., & Haier, R. J. (2007). The parieto-frontal integration theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences, 30*, 135-187. doi: 10.1017/S0140525X07001185

K

Kane, M.J., Hambrick, D.Z., Tuholski, S.W., Wilhelm, O., Payne, T.W., & Engle, R.W. (2004). The generality of working memory capacity: a latent variable approach to verbal end visuospatial memory span and reasoning. *Journal of Experimental Psychology: general, 133,* 189-217. doi: 10.1037/0096-3445.133.2.189

Kaufman, A.S. (2000). Tests of Intelligence. In: R.J. Sternberg (Ed), *Handbook of Intelligence* (pp. 445-476). New York: Cambridge University Press.

Kaufman, A. S., & Kaufman, N. L. (1993). *Manual for the Kaufman Adolescent and Adult Intelligence Test (KAIT)*. Circle Pines, MN: American Guidance Service.

Kaufman, J.C., Kaufman, S.B., & Plucker, J.A. (2013). Contemporary theories of intelligence. In J. Reisberg (Ed.) *Oxford Handbook of Cognitive Psychology* (pp. 811-822). New York: Oxford University Press.

Kaufman, J.C., Lichtenberger, E.O., & Kaufman, A.S. (2003). Assessing the intelligence of adolescents with the Kaufman Adolescent and Adult Intelligence test (KAIT). In: C.R. Reynolds & R.W. Kamphaus (Eds.). *Handbook of psychological and educational assessment of children: Intelligence, aptitude, and achievement* (2nd ed., pp. 174-186). New York: Guilford Press.

Kaufman, A.S., & Lichtenberger, E.O. (1999). *Essentials of WAIS-III assessment.* New York: John Wiley & Sons, Inc.

Kaufman, A.S., Lichtenberger, E.O., & McLean, J.E. (2001). Two- and three-factor solutions of the WAIS-III. *Psychological Assessment, 8,* 267-280. doi: 10.1177/107319110100800303

Keith, T.Z., & Reynolds, M.R. (2010). Cattell-Horn-Carroll abilities and cognitive tests: what we've learned from 20 years of research. *Psychology in the Schools, 47,* 635-650. doi: 10.1002/pits.20496

Kooij, A.P. & Dek, J.E. (2012). *WAIS-IV-NL Technisch Rapport (The Dutch language version of the WAIS-IV: technical report).* Amsterdam, The Netherlands: Pearson Assessment and Information B.V.

Krabbendam, L., & Kalff, A. (1997). *Behavioural Assessment of the Dysexecutive Syndrome. Handleiding.* Amsterdam: Pearson Assessment and Information B.V.

L

Lezak, M.D., Howieson, D.B., Bigler, E.D., & Tranel, D. (2012). *Neuropsychological Assessment (5th ed.).* New York: Oxford University Press

Lezak, M.D., Howieson, D.G., & Loring, D.W. (2004). *Neuropsychological Assessment* (4th ed.). New York: Oxford University Press.

Lichtenberger, E.O., & Kaufman, A.S. (2009). *Essentials of WAIS-IV Assessment*. New York, NY: Wiley.

Lowe, C., & Rabbit, P. (1998). Test/ re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: theoretical and practical issues. *Neuropsychologia, 36*, 915 - 923.

Luria, A.R. (1980). *Higher Cortical Functions in man* (2nd ed). New York: Basic Books.

M

Martin, T.A., Donders, J., & Thompson, E. (2000). Potential of and problems with new measures of psychometric intelligence after traumatic brain injury. *Rehabilitation Psychology, 45,* 402-408. doi: 10.1037/0090-5550.45.4.402

McCabe, D.P., Roediger III, H.L., McDaniel, M.A., Balota, D.A., & Hambrick, D.Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology, 24,* 222-243. doi: 10.1037/a0017619

McFarland, D.J. (2013). Modeling individual subtests of the WAIS-IV with multiple latent factors. *PLoSONE, 8,* e74980. doi: 10.1371/journal.pone.0074980

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37,* 1-10. doi: 10.1016/j.intell.2008.08.004

McGrew, K.S (2010). New directions in neuropsychological assessment: Augmenting neuropsychological assessment with CHC cognitive measures. The Australian Psychological Society, *16th Annual APS College of Clinical Neuropsychologists Conference From East to West: New directions in Neuropsychology.* Retrieved from http://www.iqscorner.com/2010/10/new-directions-in-neuropsychological.html

McGurk, S.R., Coleman, T., Harvey, P.D., Reichenberg, A., White, L., Friedman, J., …& Davis, K.L. (2004). Working memory performance in poor outcome schizophrenia: Relationship to age and executive functioning. *Journal of Clinical and Experimental Neuropsychology, 26,* 153-160. doi: 10.1076/jcen.26.2.153.28079

Miyake, A., Friedman, P. F., Emerson, M.J., Witzki, A.H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: a latent variable analysis. *Cognitive Psychology, 41*, 49-100. doi: 10.1006/cogp.1999.0734

Mulder, J.L., Dekker, R., & Dekker, P.H. (2005). *Kaufman – Intelligentietest voor Adolescenten en Volwassenen.* Leiden: PITS.

Müller, V. I., Langner, R., Cieslik, E. C., Rottschy, C., & Eickhoff, S. B. (2015). Interindividual differences in cognitive flexibility: influence of gray matter volume, functional connectivity and trait impulsivity. Brain Structure and Function, 220(4), 2401-2414. doi: 10.1007/s00429-014-0797-6

N

Naglieri, J.A., & Goldstein, S. (2013). *Comprehensive executive functioning inventory technical manual.* Toronto, Canada: Multi-Health Systems.

Niileksela, C. R., Reynolds, M. R., & Kaufman, A. S. (2013). An alternative Cattell-Horn- Carroll (CHC) factor structure of the WAIS-IV: Age invariance of an alternative model for ages 70-90. *Psychological Assessment, 25,* 391-404. doi: 10.1037/a0031175

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R.J. Davidson, G.E. Schwartz, & D. Shapiro (Eds.). *Consciousness and self-regulation: Advances in research and theory* (vol. 4, pp. 1–18). New York: Plenum.

Nyhus, E., & Barceló, F. (2009). The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: A critical update. *Brain and Cognition, 71,* 437-451. doi: 10.1016/j.bandc.2009.03.005

O

Oosterman, J.M., & Scherder, E.J.A. (2006). Distinguishing between vascular dementia and Alzheimer's disease by means of the WAIS: A meta-analysis. *Journal of Clinical and Experimental Neuropsychology, 28,* 1158-1175. doi: 10.1080/13803390500263543

Oosterman, J.M., Vogels, R.L.C., Van Harten, B., Gouw, A.A., Poggesi, A., Scheltens, P., Kessels, R.P.C., & Scherder, E.J.A. (2010). Assessing mental flexibility: neuroanatomical and neuropsychological correlates of the Trail Making Test in elderly people. *The Clinical Neuropsychologist, 24,* 203-129. doi: 10.1080/13854040903482848

Owen, A.M., Downes, J.J., Sahakian, B.J., Polkey, C.E., & Robbins, T.W. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia, 28*, 1021-1034. doi: 10.1016/0028-3932(90)90137-D

P

Packwood, S., Hodgetts, H. M., & Tremblay, S. (2011) A multiperspective approach to the conceptualization of executive functions. *Journal of Clinical and Experimental Neuropsychology, 33*, 456-470. doi:10.1080/13803395.2010.533157

Parsey, C.M., & Schmitter-Edgecombe, M. (2013). Applications of technology in neuropsychological assessment. *The Clinical Neuropsychologist, 27,* 1328-1361. doi: 10.1080/13854046.2013.834971

Piaget, J. (2008). Intellectual evolution from adolescence to adulthood. *Human Development, 51*, 40-47. doi: 10.1159/000112531

R

Redick, T.S., Unsworth, N., Kelly, A.J., & Engle, R.W. (2012). Faster, smarter? Working memory capacity and perceptual speed in relation to fluid intelligence. *Journal of Cognitive Psychology, 24,* 844-854. doi: 10.1080/20445911.2012.704359

Resing, W.C.M. (2016). Dynamisch en statisch testen: Onmisbaar in psychodiagnostisch onderzoek. *Tijdschrift voor Neuropsychologie, 1,* 52-64.

Robbins, T.W., James, M., Owen, A.M., Sahakian, B.J., Lawrence, A.D., McInnes, J., & Rabbitt, P.M.A. (1998). A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: Implications for theories of executive functioning and cognitive ageing. *Journal of the International Neuropsychological Society, 4*(5), 474-490. Retrieved from http://journals.cambridge.org/download.php?file=%2FINS%2FINS4_05%2FS135567798455073a.pdf&code=a3cd65eacab83307cc1b3917051f9ad5

Roberds, E.L. (2015). *Evaluation the relationship between CHC factors and executive functioning.* (Doctoral dissertation). Retrieved from http://cardinalscholar.bsu.edu/handle/123456789/199524

Roca, M., Manes, F., Chade, A., Gleichgerrcht, E., Gershanik, O., Arévalo, G.G., … Duncan, J. (2012). The relationship between executive functions and fluid intelligence in Parkinson's disease. *Psychological Medicin, 22* 1-8. doi: 10.1017/S0033291712000451

Roca, M., Manes, F., Gleichgerrcht, E., Watson, P., Ibanez, A., Thompson, R., Torralva, T., & Duncan, J. (2013). Intelligence and executive functions in frontotemporal dementia. *Neuropsychology, 51,* 725-730. doi: 10.1016/j.neuropsychologia.2013.01.008

Roca, M., Manes, F., Cetkovich, M., Bruno, D., Ibanez, A., Torralva, T., & Duncan, J. (2014). The relationship between executive functions and fluid intelligence in schizophrenia. *Frontiers in Behavioral Neuroscience, 8,* 46. doi: 10.3389/fnbeh.2014.00046

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., … Duncan, J. (2010). Executive function and fluid intelligence after frontal lobe lesions. *Brain, 133*, 234-247. doi: 10.1093/brain/awp269

Ryan, J.J. & Paolo A.M., (2001). Exploratory factor analysis of the WAIS–III in a mixed patient sample. *Archives of Clinical Neuropsychology, 16,* 151-156. doi: http://dx.doi.org/10.1016/S0887-6177%2899%2900061-X

S

Salthouse, T.A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology, 19,* 532-545. doi: 10.1037/0894-4105.19.4.532

Salthouse, T.A., Atkinson, T.M., & Berish, D.E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General, 132*, 566-594. doi: 10.1037/0096-3445.132.4.566

Salthouse, T.A., & Davis, H.P. (2006). Organization of cognitive abilities and neuropsychological variables across the lifespan. *Developmental Review, 26*, 31-54. doi: 10.1016/j.dr.2005.09.001

Salthouse, T.A., & Pink, J.E. (2008). Why is working memory related to fluid intelligence? *Psychonomic Bulletin & Review, 15*, 364-371. doi: 10.3758/PBR.15.2.364

Sánchez-Cubillo, I., Periánez, J.A., Adrover-Roig, D., Rodríguez-Sánchez, J.M., Ríos-Lago, M., Tirapu, J., & Barceló. (2009). Construct validity of the Trail Making Test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *Journal of the International Neuropsychological Society, 15,* 438-450. doi: 10.1017/S1355617709090626

Satorra, A. (1992). Asymptotic robust inferences in the analysis of mean and covariance structures. *Sociological Methodology, 22,* 249-278. Retrieved from http://www.econ.upf.edu/~satorra/CourseSEMVienna2010/SatorraSM1992.pdf

Satorra, A., & Bentler, P.M. (1990). Model conditions for asymptotic robustness in the analysis of linear relations. *Computational Statistics & Data Analysis, 10,* 235-249. doi: http://dx.doi.org/10.1016/0167-9473(90)90004-2

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. Flanagan & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99-144). New York, NY: Guilford.

Schreiber, J.B., Nora, A., Stage, F.K., Barlow, E.A. & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: a review. *The Journal of Educational Research, 99,* 323-338. doi: 10.3200/JOER.99.6.323-338

Schwartz, D. M. (2013). A four- and five- factor structural model for Wechsler tests: Does it really matter clinically? *Journal of Psychoeducational Assessment, 31,* 175-185. doi: 10.1177/0734282913478038

Schweizer, K., Troche, S.J., & Rammsayer, T.H. (2011). On the special relationship between fluid and general intelligence: new evidence obtained by considering the position effect. *Personality and Individual Differences, 50,* 1249-1254. doi: 10.1016/j.paid.2011.02.019

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London, 298,* 199-209. doi: 10.1098/rstb.1982.0082

Spearman, C. (1927). *The abilities of man.* Oxford, England: Macmillan SPSS Inc. Released 2009. PASW Statistics for Windows, Version 18.0. Chicago: SPSS Inc.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 18, 643-662. doi: 10.1037/h0054651

Sweet, L.H., Paskavitz, J.F., O'Conner, M.J., Bromndyke, J.N., Wellen, J.W., & Cohen, R.A. (2005). fMRI correlates of the WAIS-III Symbol Search subtest. *Journal of the International Neuropsychological Society, 11,* 471-476.

T

Taylor, M.J., & Heaton, R.K. (2001). Sensitivity and specificity of WAIS-III/WMS-III demographically corrected factor scores in neuropsychological assessment. *Journal of the International Neuropsychological Society, 7,* 867-874.

Taub, G.E. (2001). A confirmatory analysis of the Wechsler Adult Intelligence Scale- third edition: is the verbal/ performance discrepancy justified? *Practical Assessment, Research & Evaluation, 7*

Thompson, B. (2000). Ten commandments of structural equation modeling. In L.G. Grimm P.R. Yarnold (Eds.), *Reading and understanding MORE multivariate statistics* (pp. 261-283). Washington, DC: American Psychological Association.

U

Unsworth, N., & Engle, R.W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory & Language, 54*, 68-80. doi: 10.1016/j.jml.2005.06.003

Unterrainer, J.M., Rahm, B., Kaller, C.P., Leonhart, R., Quiske, K., Hoppe-Seyler, K., … Halsband, U. (2010). Planning abilities and the Tower of Londen: Is this task measuring a discrete cognitive function? *Journal of Clinical and Experimental Neuropsychology, 26,* 846-856. doi: 10.1080/13803390490509574

V

Van Aken, L., Kessels, R.P.C., Wingbermühle, E., Van der Veld, W.M., & Egger, J.I.M. (2015a). Fluid intelligence and executive functioning more alike than different? *Acta Neuropsychiatrica, 28,* 31-37. doi: 10.1017/neu.2015.46

Van Aken, L., Kessels, R.P.C., Wingbermühle, E., Wiltink, M., Van der Heijden, P.T., Egger, J.I.M. (2014). Exploring the incorporation of executive functions in intelligence testing: factor analysis of the WAIS-III and traditional tasks of executive functioning. *International Journal of Applied Psychology, 4,* 73-80. doi: 10.5923/j.ijap.20140402.05

Van Aken, L., Van der Heijden, P.T., Van der Veld, W.M., Hermans, L., Kessels, R.P.C., & Egger, J.I.M. (2015b). Representation of the Cattell-Horn-Carroll Theory of Cognitive Abilities in the Factor Structure of the Dutch-Language Version of the WAIS-IV. *Assessment,* 1-9. doi: 10.1177/1073191115607973

Van der Elst, W., Van Boxtel, M.P.J., Van Breukelen, G.J.P., & Jolles, J. (2006). Normative data for the Animal, Profession and letter *M* Naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex. *Journal of the International Neuropsychological Society, 12,* 80-89. doi: 10.1017/S1355617706060115

Van der Elst, W., Van Boxtel, M.P.J., Van Breukelen, G.J.P., & Jolles, J. (2005). Rey's verbal learning test: Normative data for 1855 healthy participants aged 24-81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society, 1,* 290-302. doi: 10.1017/S1355617705050344

Van der Elst, W., Van Boxtel, M.P.J., Van Breukelen, G.J.P., & Jolles, J. (2006). The Stroop Color-Word Test. Influence of Age, Sex, and Education; and Normative Data for a Large Sample Across the Adult Age Range. *Assessment, 13,* 62-79. doi: 10.1177/1073191105283427

Van der Heijden, P., & Donders, J. (2003a). A Confirmatory factor analysis of the WAIS-III in patients with traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 25*, 59-65. doi: 10.1076/jcen.25.1.59.13627

Van der Heijden, P., & Donders, J. (2003b). WAIS-III factor index score patterns after traumatic brain injury. *Assessment, 10,* 115-122. doi: 10.1177/1073191103010002001

Van der Heijden, P., Van den Bos, P., Mol, B., & Kessels, R.P.C. (2013). Structural Validity of the Dutch-Language Version of the WAIS-III in a Psychiatric Sample. *Applied Neuropsychology: Adult, 20,* 41-46. doi: 10.1080/09084282.2012.670166

Van der Meer, J.M.J., & Eling, P. (2008). Andre Rey en zijn Complexe Figuur Test. *Tijdschrift voor Neuropsychologie, 3,* 45-51.

Van Ravenzwaaij, D., & Hamel, R. (2006). De Nederlandstalige WAIS-III na hernormering [The Dutch WAIS-III after revision]. *De Psycholoog, 41,* 268 – 271.

W

Ward, L. C., Bergman, M. A., & Hebert, K. R. (2012). WAIS-IV subtest covariance structure: conceptual and statistical considerations. *Psychological Assessment, 24,* 328-340. doi: 10.1037/a0025614

Wechsler, D. (1997a). *Technical manual for the Wechsler Adult Intelligence Test – third edition*. San Antonio: The Psychological Corporation.

Wechsler, D. (1997b). *WAIS-III: Wechsler Adult Intelligence Scale - third edition: Administration and scoring manual*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2000). *WAIS-III Nederlandstalige bewerking: Afname en scorehandleiding* [*WAIS-III: Administration and scoring manual*]. Lisse: Swets & Zeitlinger.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale - fourth edition*. San Antonio, TX: Pearson.

Wechsler, D. (2012a). WAIS-IV-NL *Nederlandstalige bewerking, technische handleiding (The Dutch language version of the WAIS-IV: technical manual).* Amsterdam, The Netherlands: Pearson Assessment and Information B.V.

Wechsler, D. (2012b). WAIS-IV-NL *Nederlandstalige bewerking, afname- en scoringshandleiding (The Dutch language version of the WAIS-IV: administration and scoring manual)*. Amsterdam, The Netherlands: Pearson Assessment and Information B.V.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children – Fifth Edition*. San Antonio, TX: Pearson.

Weiss, L. G., Keith, T. Z., & Zhu, J., & Chen, H. (2013a). WAIS-IV and clinical validation of the four- and five-factor interpretative approaches. *Journal of Psychoeducational Assessment, 31,*94-113. doi: 10.1177/0734282913478030

Weiss, L. G., Keith, T. Z., & Zhu, J., & Chen, H. (2013b). Technical and practical issues in the structure and clinical invariance of the Wechsler Scales: a rejoinder to commentaries. *Journal of Psychoeducational Assessment, 31,*235-243. doi: 10.1177/0734282913478050

Williams, B., Myerson, J., & Hale, S. (2008). Individual differences, intelligence, and behaviour analysis. *Journal of the experimental analysis of behaviour, 90*, 219-231. doi: 10.1901/jeab.2008.90-219

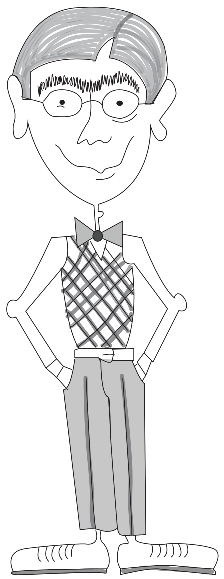
Wilson, B.A., Alderman, N., Burgess, P.W., Emslie, H., & Evans, J.J. (1996). *Behavioural Assessment of the Dysexecutive Syndrome (BADS).* St. Edmunds, UK: Thames Valley Test Company.

Wood, R.L., & Liossi, C. (2007). The relationship between general intellectual ability and performance on ecologically valid executive tests in a severe brain injury sample. *Journal of the International Neuropsychological Society, 13,* 90-98. doi: 10.10170S1355617707070129

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson III. Itasca,IL. Riverside.

Woolgar, A., Parr, A., Cusack, R., Thompson, R., Nimmo-Smith, I., Torralva, T., ... Duncan, J. (2010). Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 14899-14902. doi: 10.1073/pnas.1007928107

Nederlandse samenvatting



Nederlandse samenvatting

Doel van dit proefschrift is om meer inzicht te krijgen in de relatie tussen executief functioneren (EF) en intelligentie, in het bijzonder het Cattell-Horn-Carroll model van intelligentie. EF is een begrip dat veel gebruikt wordt binnen de neuropsychologische praktijk en verwijst naar verschillende cognitieve vaardigheden zoals het plannen, monitoren, aanpassen en inhiberen van gedrag. Deze vaardigheden hangen samen met, maar zijn ook te onderscheiden van intelligentie. Binnen de theorievorming over intelligentie is het CHC-model een van de meest invloedrijke theorieën. Dit model onderscheid verschillende vormen van intelligentie zoals gekristalliseerde kennis en vloeiende capaciteiten, maar EF wordt hierin niet expliciet beschreven. Dit is opmerkelijk, omdat EF binnen de neuropsychologische praktijk een belangrijke plaats heeft en beide benaderingen conceptueel overlap vertonen binnen het domein van de cognitieve vaardigheden. In onderhavige dissertatie wordt de relatie tussen EF en intelligentie uitgediept op basis van metingen met verschillende instrumenten in diverse klinische steekproeven. Het doel is om het klinisch relevante concept EF te verbinden met een factoranalytisch model over intelligentie. In dit hoofdstuk wordt een samenvatting gegeven van de voornaamste bevindingen van de studies en de overwegingen daarbij, als ook van de sterktes en zwaktes van de gebruikte methoden en technieken. Tot slot worden implicaties voor toekomstig onderzoek besproken.

Samenvatting van de studies

In **hoofdstuk 2** wordt de representatie van EF in de derde editie van de Wechsler Adult Intelligence Test (WAIS-III; Wechsler, 1997) onderzocht. De Wechsler intelligentie-schalen zijn de meest gebruikte intelligentietests in de praktijk, maar er is nog onduidelijk-heid over de wijze waarop je EF met de WAIS-III kunt meten. In de WAIS-III worden scores van 14 subtests samengevat in een verbale (VIQ) en performale (PIQ) component, welke samen een totaal IQ (TIQ) vormen. Een onderverdeling in vier factoren bleek zowel vanuit theoretisch als vanuit klinisch perspectief geschikter voor de indeling van de subtests dan de VIQ – PIQ-dichotomie (Kaufman & Lichtenberger, 1999). Deze vier factoren zijn de Verbale Begripsindex (VBI), de Perceptuele Organisatie-index (POI), de Werkgeheugenindex (WGI) en de Verwerkingssnelheidsindex (VSI).

Hoewel in de ontwikkeling van de WAIS-III vooral empirische en pragmatische argumenten een rol hebben gespeeld, bleken de WAIS-III-indexen redelijk goede representaties van vijf theoretische constructen uit het CHC-model te zijn, te weten gekristalliseerde intelligentie (Gc), vloeiende intelligentie (Gf), visuele informatie-verwerking (Gv), korte termijn- en werkgeheugen (Gsm) en Verwerkingssnelheid (Gs) (Alfonso, Flanagan & Radwan, 2005). Tegelijkertijd blijkt de WAIS-III voornamelijk gekristalliseerde vaardigheden te meten, terwijl het aandeel van andere vaardigheden, waaronder Gf, beperkt is (Blair, 2006; Duncan et al., 1995). Gf is kortgezegd het vermogen om nieuwe, complexe problemen op te lossen zonder daarbij een beroep te kunnen doen op aangeleerde kennis. Gf is in de literatuur veelvuldig gerelateerd aan EF. Voor het beschrijven van informatieverwerkingsprocessen is EF een belangrijk concept in de neuropsychologie en het meten van EF is dan ook een belangrijk onderdeel van het neuropsychologisch onderzoek.

Om het aandeel van EF in de WAIS-III te onderzoeken, werd in hoofdstuk 2 gebruik gemaakt van drie veelgebruikte, traditionele executieve taken; de BADS (Behavioural Assessment of the Dysexecutive Syndrome; Wilson, Alderman, Burgess, Emslie & Evans, 1996; Krabbendam & Kalff, 1997), de Stroop Kleur-Woord Test (Hammes, 1971), en de Wisconsin Card Sorting Test (WCST; Heaton, 1981). Middels exploratieve factoranalyse met maximum-likelihoodprocedures werd een twee-, drie- en vier-factorenmodel onderzocht. Zowel het drie- als vier-factorenmodel bleken passend bij de empirische gegevens. Het vier-factorenmodel paste het beste en kwam vrijwel volledig overeen met de factorstructuur volgens de vier indexen van de WAIS-III. Alle drie de EF-taken laadden op de factor die overeen kwam met de POI. Deze resultaten onderstrepen eerdere bevindingen waarbij een grote overlap tussen Gf en EF is aangetoond, aangezien enkele POI-taken geassocieerd zijn met Gf (Duncan, 2010; Roca et al., 2010). Verderop in dit deze samenvatting zal de relatie tussen Gf en EF nader worden besproken. De resultaten van hoofdstuk 2 bevestigen voorts het grote aandeel van Gc in de WAIS-III. Gc wordt gemeten door de VBI. Deze index verklaart maar liefst 41,7 % van het model, ten opzichte van 12,9 % verklaarde variantie door de POI, 6,2 % door de WGI en 5,2% door de VSI. Dit geeft een vertekend beeld van intelligentie waar mogelijk niet iedere clinicus zich bewust van is (Blair, 2006). De onevenredige verdeling roept tevens vragen op over de inhoudsvaliditeit van de WAIS-III, zeker wanneer men zich bedenkt dat in veel gevallen Gf een betere representatie van *g* blijkt te zijn dan Gc (Duncan, 1995, 2010).

In **hoofdstuk 3** is de meest recente versie van de Wechsler intelligentie schalen, de WAIS-IV, onderzocht. In de ontwikkeling van deze laatste versie heeft het CHC-model een leidende rol gespeeld, waarbij het beter meten van Gf een belangrijk doel was. Dezelfde vijf CHC-vaardigheden in de WAIS-III worden ook gemeten in de WAIS-IV (Benson et al., 2010; Weiss et al., 2013a). De factorstructuur is ook grotendeels vergelijkbaar met die van de WAIS-III, behalve in het geval van de POI, welke is vervangen door de Perceptuele Redeneerindex (PRI). Subtests van de POI zijn aangepast ter verbetering van het meten van vloeiende processen en de aangepaste naam sluit hier volgens de makers meer op aan (Wechsler, 2008)

Met behulp van confirmatieve factoranalyse werd een vijf-factorstructuur van de WAIS-IV gebaseerd op de vijf CHC-vaardigheden onderzocht, als alternatief op het vier-factorenmodel waar de indexscores op gebaseerd zijn. Achterliggend idee hierbij is dat de PRI, anders dan de andere factoren, geen ‘CHC-zuivere’ factor is, maar een combinatie van zowel Gf als Gv. In het alternatieve vijf factor model is de PRI opgedeeld in een Gv factor bestaande uit de subtests Blokpatronen, Matrix Redeneren, Figuur Leggen, Gewichten en Onvolledige tekeningen en een Gf factor bestaande uit de subtests Matrix Redeneren, Gewichten en Rekenen. Zowel het vier-factorenmodel volgens de indexen als het alternatieve vijf-factoren-CHC-model bleken een goede fit met de gegevens te hebben. Ondanks het principe van zuinigheid werd besloten voorkeur te geven aan het vijf-factorenmodel, omdat dit model opgebouwd is vanuit een theoretisch kader (het CHC-model). Los van de PRI correspondeerden de drie andere indices direct met de veronderstelde gemeten CHC-vaardigheden; de VBI, WGI en PSI bleken adequate representaties van respectievelijk Gc, Gsm en Gs. Een opvallend resultaat was dat het aandeel van Gf in de WAIS-IV opmerkelijk laag was. Gf werd gerepresenteerd door drie subtests, maar slechts twee van deze subtests zijn onderdeel van de kernbatterij. De bijdrage aan de PRI en het TIQ is dan ook beperkt. De poging om meer vloeiende processen in de WAIS-IV te includeren is onvoldoende waargemaakt. Desalniettemin is het huidige vijf-factorenmodel inclusief de factor Gf valide, waarbij het uit elkaar halen van de factoren Gf en Gv bijdraagt aan een beter begrip van test prestaties op de betreffende subtests.

In **hoofdstuk 4** staan twee constructen die aan de basis liggen van het CHC-model centraal; Gf en Gc. Beide constructen zijn, meer dan de andere CHC-vaardigheden, nauw verwant aan elkaar én aan *g* (Benson et al., 2010; Horn & Cattell, 1966; Grégoire, 2013; Schweizer, Troche & Rammsayer, 2011; Schneider & McGrew, 2012). In dit hoofdstuk wordt een neuropsychologisch perspectief gehanteerd, waarbij de veronderstelde relatie tussen Gf en EF nader wordt onderzocht. Op het eerste gezicht vertonen definities van EF en Gf veel overeenkomsten. Beiden beschrijven een efficiëntie in handelen bij de uitvoering van complex gedrag van niet-geautomatiseerde gedragingen in nieuwe of onbekende situaties. Overlap tussen beide constructen is al eerder aangetoond en resultaten uit hoofdstuk 2 suggereren eveneens een relatie tussen EF en Gf. Zoals blijkt uit hoofdstuk 2 en 3 is de representatie van Gf in de Wechslerschalen beperkt. Daarom is nu de Kaufman Adolescent and Adult Intelligence Test (KAIT) gebruikt als maat voor zowel Gf als Gc, waarbij een latent model is ontwikkeld om de relaties tussen EF, Gf en Gc te meten. Drie taken van de Cambridge Automated Neuropsychological Test Battery (CANTAB) zijn geselecteerd als maten voor EF, waarbij onder andere het werkgeheugen, planningsvermogens en mentale flexibiliteit zijn gemeten. De resultaten laten een sterke samenhang zien tussen EF en Gf (correlatie van 0,91), anders dan de samenhang tussen EF en Gc, en Gf en Gc onderling. Werkgeheugen lijkt een cruciale rol te spelen in deze relatie, wat te zien is aan een hoge bijdrage van de werkgeheugentest die wordt geleverd aan de EF-Gf-relatie.

De huidige resultaten komen overeen met eerder onderzoek naar de EF-Gf-relatie. De bijzonder hoge correlatie die werd gevonden lijkt deels gevolg van de geselecteerde taken. Gf kan worden beschouwd als een algemene, domeinonafhankelijke vaardigheid (Duncan, 2012). Er bestaat een grote variatie in EF-taken, waarbij de CANTAB-taken als complex en breed van opzet worden beschouwd in vergelijking met relatief enkelvoudige reactietijdtaken of eenvoudige *shifting*- of inhibitietaken zoals een *go/no-go*-paradigma. Dit zou de hoge correlatie met Gf kunnen verklaren.

Diamond (2013) en Duncan (2010, 2012) stellen dat inadequate taakuitvoering niet alleen van specifieke vaardigheden afhankelijk is, maar ook bepaald worden door de beoordeling van bepaalde taakvereisten (hoe worden taakvereisten gerangschikt naar belangrijkheid) en door de interactie tussen deze vereisten om zo tot doelgericht gedrag te komen. Deze overwegingen in acht nemend, wordt geconcludeerd dat er onderscheid gemaakt moet worden tussen specifieke versus algemene cognitieve processen bij het meten van EF.

Intelligentieonderzoek is tegenwoordig vaak een standaardonderdeel van neuropsychologisch onderzoek. In **hoofdstuk 5** wordt de relatie tussen veelgebruikte executieve taken en intelligentiematen (als representatie van CHC-constructen) onderzocht. In de eerdere studies werd factoranalyse gebruikt om zowel EF- als CHC-vaardigheden te onderzoeken. Waar de focus dus eerder op latente constructen lag, gaat hoofdstuk 5 in op de operationalisatie van deze constructen en hun bruikbaarheid op gedragsniveau.

Met behulp van multivariate regressieanalyses werd de voorspellende waarde van EF-taken op CHC-domeinen gemeten. EF-taken bleken gemiddeld 50% van de variantie in de intelligentietests te verklaren. Dit komt overeen met eerder factoranalytisch onderzoek. Er werden relatief lage correlaties tussen EF-taken onderling gevonden, terwijl correlaties tussen CHC-maten onderling hoger waren. Laatstgenoemde kan erop duiden dat CHC-maten 'verzadigd' zijn met *g* en derhalve weinig onderscheidend vermogen hebben.

De Tower of London (ToL; planningstaak) en de Stroop Kleur-Woord Test (inhibitietaak) bleken CHC-vaardigheden het minst te voorspellen. Dit impliceert dat planningsvaardigheden en inhibitoire processen beperkt worden gemeten met de huidige intelligentietests. Een andere verklaring is dat de ToL en de Stroop over onvoldoende psychometrische kwaliteiten beschikken. Hier draagt ook de beperkte constructvaliditeit van voornamelijk de ToL aan bij. Deze taakonzuiverheid van de ToL wordt ook bij andere executieve taken gezien (Miyake et al., 2000; Salthouse, 2005) en vormt een fundamenteel probleem in het adequaat meten van executieve processen.

Concluderend wijzen resultaten er op dat EF-CHC-relaties zoals die zijn bestudeerd op een latent niveau niet direct zijn te vertalen naar gedragsniveau middels het gebruik van manifeste variabelen, oftewel instrumenten die standaardonderdeel zijn van het neuropsychologisch onderzoek. Dit is het gevolg van moeilijkheden bij de operationalisatie van zowel EF als intelligentieconstructen.

Bovenstaande resultaten kunnen in de volgende drie punten worden samengevat. **Ten eerste**, er is sprake van een grote overlap tussen EF en intelligentie. Voornamelijk Gf lijkt processen te behelzen die overeenkomen met EF. Hoewel de twee constructen overlappen, lijkt deze overlap mede afhankelijk van de complexiteit van de gebruikte instrumenten. **Ten tweede**, de neuropsychologische benadering van EF en de psychometrische CHC-hiërarchie zijn beide succesvol in het beschrijven van cognitief functioneren, maar beide vertonen ook hiaten. De neuropsychologie conceptualiseert adequaat (executieve) processen van informatieverwerking, waarbij echter de vertaling van deze theorieën naar bruikbare instrumenten tot op heden beperkt blijkt. Het CHC-model resulteert in relatief zuivere metingen van latente constructen, maar de vertaling naar hoe deze constructen zich manifesteren en verhouden tot informatieverwerkingsprocessen in het brein, is nog onvoldoende duidelijk gemaakt. Ook is het nog de vraag hoe de onderliggende CHC-factoren samenhangen met het functioneren in de dagelijkse praktijk. **Tot slot**, er bestaat een kloof tussen bekende theorieën over EF en intelligentie enerzijds en de operationalisatie in de klinische praktijk anderzijds. In de interpretatie van (neuro)psychologisch testonderzoek moeten zowel specifieke als algemene cognitieve vaardigheden worden onderscheiden, waarbij de overlap tussen taken in acht wordt genomen. Nu beschrijven neuropsychologen nog vaak een gefragmenteerd beeld van het executief functioneren, waarbij algemene cognitieve vaardigheden buiten beschouwing worden gelaten. Wat intelligentietests betreft, ontbreekt er een duidelijke link naar de beschrijving van cognitieve processen. Executieve vaardigheden worden vooralsnog onvoldoende gemeten door de huidige intelligentietests.

In de volgende sectie worden enkele resultaten uitgelicht. De sterke relatie tussen Gf en EF wordt besproken, als ook hoe de neuropsychologische en psychometrische benaderingen van elkaar kunnen profiteren. Daarnaast zullen implicaties van bovenstaande resultaten voor de klinische praktijk worden bediscussieerd.

Algemene discussie

De Gf – EF relatie

Het wordt breed aanvaard dat Gf de kern is van complex en intelligent gedrag. Echter, ondanks de prominente plek die Gf heeft in het CHC-model en de sterke overlap met EF die gevonden wordt in hoofdstuk 4, is het aandeel van Gf in de Wechslerschalen beperkt (zie hoofdstuk 2 en 3).

Gf is zowel aan executieve processen als aan *g* gerelateerd. Het CHC-model beschrijft Gf als een algemene vaardigheid die, onafhankelijk van de cognitieve domeinen die tijdens een handeling worden aangesproken (Schneider & McGrew, 2012), betrokken is bij nieuwe en complexe situaties waar aangeleerde en automatische responsen niet volstaan. Een vergelijkbare beschrijving wordt gegeven door Duncan (2010; 2012) wanneer hij schrijft over het *Multiple Demand System* (MDS; zie de algemene inleiding voor een toelichting). In het MDS wordt Gf gezien als de *efficiëntie* waarmee men zich een nieuwe of complexe taak eigen maakt, los van het type taak en de specifieke (executieve) vaardigheden die worden aangesproken tijdens de uitvoering ervan. Het MDS vertoont dus samenhang met EF, omdat executieve processen ook betrokken zijn bij taken die gecontroleerde responsen vragen (in tegenstelling tot automatisch handelen). De bemoeienis van het MDS met de uitvoer van taken wordt groter bij complexe taken. Hoe complexer een EF taak is opgebouwd, hoe meer de betrokkenheid van het MDS, dus hoe groter de overlap met Gf.

De zienswijzen van zowel Schneider e McGrew (2012) als Duncan (2010; 2012) impliceren dat elk type cognitieve test een Gf-component bezit die groter is al naar gelang de complexiteit en/of nieuwheid van de test. De moeilijkheid zit in het identificeren van deze vloeiende component. Want hoewel Gf als 'domeinonafhankelijk' (dus los van domeinen als geheugen, aandacht, taal of motoriek) kan worden beschouwd, is het onmogelijk om een taak te ontwikkelen die alleen maar vloeiende vaardigheden aanspreekt (Grégoire, 2013). Elke taak bestaat dus uit algemene vloeiende vereisten (namelijk de efficiëntie om met complexiteit om te gaan), naast specifieke cognitieve vereisten zoals het *switchen* van de aandacht of impulsen onderdrukken. Een logisch gevolg is dan ook dat uitkomstmaten van een taak hierop ingericht worden. Nu worden taken veelal ingezet als maat van één specifieke vaardigheid, waarna prestaties op verschillende taken vervolgens met elkaar worden vergeleken. Er dient echter b*í*nnen taken meer aandacht te zijn voor de simultane cognitieve processen, waarbij onderscheid gemaakt zou moeten worden tussen algemene cognitieve vaardigheden en specifieke cognitieve vereisten. Op deze manier kan ook samenhang tussen cognitieve constructen vanuit een ander perspectief worden geïnterpreteerd.

Het CHC model en executieve functies

Neuropsychologische theorieën over EF beschrijven pathologisch gedrag in termen van cognitief disfunctioneren en bijbehorende klinisch-relevante gedragingen. Het CHC-model stamt uit de academische psychologie en streeft naar een zuivere beschrijving van alle mogelijke cognitieve vaardigheden. Idealiter smelten de twee benaderingen samen, waarbij alle mogelijke cognitieve factoren worden geïdentificeerd die een voorspellende waarde hebben voor klinisch relevante pathologische gedragingen. Zo ver is het echter nog niet, zoals moge blijken uit hoofdstuk 5. Dit komt mede door het feit dat veel tests sterk leunen op *g,* terwijl ze vaak als domeinspecifieke taken worden geïnterpreteerd. De samenhang tussen de theoretische constructen en de instrumenten die deze constructen pogen te meten is echter vaak beperkt. Met andere woorden, op dit gebied bestaat een kloof tussen theorie en praktijk. Taakonzuiverheid in de operationalisatie van EF heeft sommigen doen concluderen dat EF overbodig is in het huidige CHC-model en dat EF volledig verklaard kan worden door Gf, Gv, Gs en Gsm (Jewsbury et al., 2016). Dit is enigszins bevreemdend, omdat er in het CHC-model geen theoretische overwegingen over informatieverwerking zijn opgenomen. Anderen bekritiseren de interpretatie van intelligentietests op het niveau van de indices, aangezien de unieke verklaarde variantie van de indices bovenop (de verklaarde variantie van) het TIQ te verwaarlozen is (Canivez & Watkins, 2010; Gignac, 2008; Gignac & Watkins, 2013). In de klinisch-neuropsychologische praktijk is de interpretatie van enkel het TIQ als maat niet valide als het gaat om het beschrijven van pathologisch gedrag, bijvoorbeeld in het geval van 'disharmonische profielen' bij mensen met niet aangeboren hersenletsel.

De heldere en overzichtelijke conceptualisatie van het CHC-model vormt een verleidelijk uitgangspunt voor testontwikkeling. Echter, hoewel dit psychometrisch model uitstekend verschillende aspecten of factoren van het cognitief functioneren kan *identificeren,* is het onvoldoende *verklarend* voor het begrip van de complexiteit van (pathologisch) gedrag. In dat opzicht biedt de neuropsychologische benadering van EF een beter theoretisch kader. De vertaling van EF in valide instrumenten blijkt echter gemakkelijker gezegd dan gedaan; sommige neuropsychologische theorieën (e.g. PASS model van Luria, 1980) hebben nog niet of onvoldoende hun weg gevonden in de neuropsychologische testbatterijen. Daar komt bij dat de meest gebruikte executieve taken dateren uit de eerste helft van de twintigste eeuw, gebaseerd op verouderde of achterhaalde ideeën over individuele verschillen en cognitieve functies. Deze tests meten bijvoorbeeld alleen de eindscore van een opgelost probleem, zonder dat ze inzicht geven in het doorgelopen proces. De kloof tussen theorie en praktijk zal blijven bestaan zolang neuropsychologen enkel vasthouden aan oude tradities. Hetzelfde geldt overigens voor psychometrici en latente modellen.

Klinische implicaties en overwegingen

Het is duidelijk dat de vertaling van theoretisch construct naar een test score complex is. Dit heeft implicaties voor het meten van EF en intelligentie in de dagelijkse praktijk, als ook voor het ontwikkelen van nieuw test materiaal.

Executief disfunctioneren kan leiden tot een verlaagde POI bij de WAIS-III, resulterend in een 'disharmonisch profiel'. Een afwijkende POI geeft daarom aanleiding voor verder executief onderzoek, zeker omdat het aandeel van EF in de intelligentietest zelf klein is. Hetzelfde geldt voor de WAIS-IV, waar het kleine aandeel van Gf ervoor zorgt dat aanvullend neuropsychologisch onderzoek nodig is om cognitieve processen verder in kaart te brengen. Tegelijkertijd bestaan er geen neuropsychologische taken die in een adequate Gf meting voorzien. Gezien de essentiële rol van Gf in het verklaren van complex gedrag, is het aan te bevelen dat het aandeel van Gf in de ontwikkeling van nieuwe intelligentietests worden vergroot. Ook het opnemen van executieve processen in intelligentietests draagt bij aan de klinische bruikbaarheid en mogelijkheden tot neuropsychologische interpretatie.

In hoofdstuk 3 wordt een alternatief vijf-factormodel van de WAIS-IV gepresenteerd, gebaseerd op het CHC-model. Als clinicus moet men zich er bewust van zijn dat de PRI bestaat uit zowel Gv als Gf wanneer men de WAIS-IV volgens het CHC model interpreteert. Ook het feit dat indices met elkaar samenhangen en ook op theoretisch niveau overlap vertonen, moet in acht worden genomen bij de profielinterpretatie. Daarnaast zijn tests vaak onvolledige vertalingen van theoretische constructen. Neem bijvoorbeeld de WGI van de WAIS-IV. Hoewel dit op zichzelf een valide en betrouwbare meting van het construct werkgeheugen is (of Gsm in CHC-terminologie), is het geen *volledige* representatie van dit construct; zo is de mate van manipulatie van informatie gering en er wordt geen beroep gedaan op visuospatiële vaardigheden (Egeland, 2015).

Aan de operationalisatie van EF zitten nog andere haken en ogen. Executieve taken zijn vaak te versnipperd. Het gebruik van losse taken voor vaardigheden als *shifting*, updaten en inhiberen van informatie doet lijken alsof alle tests unieke en onafhankelijke theoretische constructen meten, waardoor de interpretatie ervan te strikt is en samenhang tussen taken over het hoofd wordt gezien.

Zoals besproken in hoofdstuk 5, hangt de intensiteit van de Gf-EF-relatie af van de complexiteit van de gebruikte taken om deze relatie mee te meten. Wanneer executieve taken heel complex zijn, is de bemoeienis van Gf (oftewel activiteit van het MD systeem) vele malen groter, resulterend in een grotere gemeten samenhang tussen Gf en EF. Hieruit kan met concluderen dat taken die specifieke cognitieve vaardigheden meten (niet betrokken bij het MD systeem) niet zo nauw zullen samenhangen met Gf. Dit klopt in theorie, echter taken zoals de WCST en de Fluency blijken voornamelijk een meting van *g,* terwijl ze veelal geïnterpreteerd worden als specifieke maten voor mentale flexibiliteit en responsgeneratie. Het is aan het neuropsychologisch vakgebied om tests te ontwikkelen waarin zowel generieke als specifieke cognitieve processen kunnen worden geïdentificeerd en gemeten. Wanneer bij deze taken ook een 'complexiteitsgradiënt' kan worden ingebouwd, kunnen generieke procesessen (of Gf) en specifieke executieve functies beter onderscheiden worden. Ook dient er gebruik gemaakt te worden van procesdiagnostiek (zie ook Resing, 2016), naast het gebruik van statische uitkomstmaten waar tegenwoordig nog veel mee wordt gewerkt.

De neuropsychologie kan profiteren van de uitgebreidheid van het CHC-model en de externe validiteit kan toenemen door het gebruik van eensgezinde naamgeving van constructen. Anderzijds kan het CHC-model uitgewerkt worden tot een geïntegreerde theorie inclusief executieve processen. Een eerste poging hiertoe is gedaan door Schneider en McGrew (2012), die een model introduceren waarin CHC-vaardigheden in termen van informatieverwerking worden toegelicht.

Samenvattend, het bestuderen van cognitieve informatieverwerking op gedragsniveau dient te worden gedaan in termen van generiek en specifiek cognitief (dis) functioneren, wat leidt tot een meer gerichte differentiële diagnostiek en meer verfijnde indicatiestelling.

Beperkingen

Een mogelijke beperking van de huidige studies is het gebruik van heterogene neuropsychiatrische groepen middels *convenience sampling.* Gemiddelde scores en subtestspreiding liggen in deze groep lager dan bij de gemiddelde bevolking. Psychopathologie, medicatie, mentale belastbaarheid, motivatie en faalangst zijn allemaal factoren die invloed hebben op testprestaties. Bij patiënten met frontale pathologie worden bijvoorbeeld hogere correlaties gevonden tussen EF en intelligentiematen (Friedman et al., 2006; Rabbit. Lowe & Shilling, 2001). Het is dus niet duidelijk in hoeverre de huidige resultaten gegeneraliseerd kunnen worden naar een gezonde populatie, of naar specifieke klinische groepen of naar de volledige leeftijdsrange. Gelijke variantie tussen gezonde en klinische groepen is echter aangetoond voor zowel de WAIS-III (Van der Heijden & Donders, 2003a) als de WAIS-IV (Weiss et al., 2013a). De gevonden factorstructuren in hoofdstuk 2, 3 en 4 zijn vergelijkbaar met de factorstructuren in de normatieve data van de tests. Ook werden er geen groepsverschillen gevonden in de verhoudingen tussen correlaties (tussen EF, Gf en Gc) in hoofdstuk 3. Deze resultaten, samen met de heterogeniteit van de populatie, lijken de robuustheid van de gevonden factorstructuren te bevestigen. Daar komt bij dat het essentieel is om deze structuren in een klinisch setting te onderzoeken, juist omdat CHC weinig is onderzocht binnen dit bereik, terwijl we het wel dagelijks toepassen door middel van intelligentie onderzoek bij patiënten. Eerste studies hiernaar zijn veelbelovend, daar er geen groepsverschillen tussen gezonde en klinische groepen worden gevonden wat betreft variantie (Jewsbury et al., 2016). In de toekomst dienen deze thema's verder onderzocht te worden, in het bijzonder door het bestuderen van de relatie tussen EF en intelligentie binnen het CHC model door gebruik te maken van andere patiëntgroepen en door het vergelijken van deze verschillende groepen.

Theorieën over intelligentie zijn veelal gebaseerd op factoranalytisch onderzoek. Deze techniek vormt tevens de basis van de huidige these. Factoranalyse is een pragmatische aanpak en het abstraheren van factoren uit grote groepen data betekent niet dat deze factoren ook daadwerkelijk onderliggende dimensies representeren. De laatste jaren lijkt het gebruik van factoranalyse te hebben geleid tot een wildgroei aan factoren, mede omdat door de gebruikte statistische technieken vaak onnodig veel factoren worden geïdentificeerd (Frazier & Youngstrom, 2007). Een groter aantal factoren draagt echter niet noodzakelijkerwijs bij aan een beter begrip van pathologisch gedrag (Blair, 2010).

Clinici zullen gebruik blijven maken van (het classificeren) van factoren zolang factoranalyse de meest gebruikte techniek blijft. Hoewel de discussie over de voor- en nadelen van factoranalyse buiten het bereik van dit proefschrift ligt, raakt het wel aan de getrokken conclusie dat we meer onderscheid moeten maken tussen generieke en specifieke cognitieve processen. Het gaat er niet alleen om dat alle mogelijke factoren in kaart worden gebracht, maar vooral *hoe* deze met elkaar samenhangen en interacteren. Alternatieve technieken naast factoranalyse zijn nodig voor een beter begrip van deze processen. Te denken valt aan regressieanalyse zoals in hoofdstuk 5 is gedaan, of andere benaderingen zoals het beschrijven van gedrag in netwerkanalyses (zie ook Borsboom & Cramer, 2013).

Tot besluit

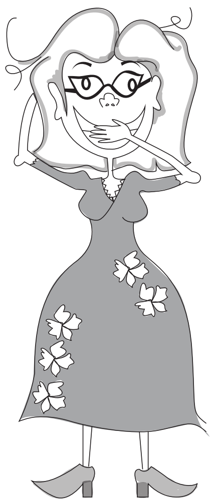
Onze klinische methodes moeten gebaseerd zijn op theoretische kaders. Psychologisch onderzoek lijkt te zijn afgedreven van het gebruik van theorieën, resulterend in een overvloed aan hypothetische constructen en een veelvoud aan neuropsychologische tests om deze te meten. Interpretatie van cognitief functioneren wordt derhalve meer bepaald door de gebruikte taken dan vanuit een theoretisch perspectief. Het lijkt alsof door de bomen het bos niet meer wordt gezien. Patiënten zijn echter gebaat bij een theoretisch denkkader, waarbij de clinicus bij de interpretatie rekening houdt met sterktes en zwaktes van de gebruikte referentiekaders. Dit zou ons moeten stimuleren om de tekorten van onze theorieën en tests te onderkennen en aan te pakken.

Zowel het CHC-model als de neuropsychologische theorieën dragen bij aan ons begrip van cognitieve functies. Over het algemeen lijkt het één niet meer of minder te verklaren dan het ander. Psychometrie is superieur in het identificeren van latente

factoren en het operationaliseren ervan. Deze latente constructen dragen echter lang niet altijd bij aan het begrip van cognitieve informatieverwerking. Cognitieve theorieën uit de neuropsychologie daarentegen, geven een duidelijke beschrijving van informatieverwerkingsprocessen. Ze doen dit wellicht te zuinig, waardoor de operationalisatie ervan tot problemen leidt. Dit beperkt het gebruik van beide benaderingen en een oplossing is niet helaas niet eenvoudig voorhanden.

De afstand tussen executieve processen en CHC-vaardigheden zal waarschijnlijk blijven bestaan en groter worden wanneer de verschillende benaderingen zich los van elkaar verder ontwikkelen. Bijvoorbeeld, psychologen werken traditioneel met lineaire technieken, terwijl het overduidelijk is dat gedrag, pathologisch gedrag bij uitstek, een niet-lineair fenomeen is. De gebruikte lineaire modellen zijn nodig om begrip te krijgen van constructen en onderliggende dimensies en hun samenhang op groepsniveau, maar dragen maar tot op zekere hoogte bij aan het begrip van afwijkend gedrag van verschillende individuen. Niet-lineaire statistische methoden en technieken kunnen behulpzaam zijn in het anders leren begrijpen van latente constructen, met meer begrip van de informatieverwerking in individuen tot gevolg.

Dankwoord



Dankwoord

Een proefschrift schrijf je (gelukkig) niet alleen. In de afgelopen jaren ben ik geholpen, gesteund en bijgestaan door velen. De meesten van jullie verdienen een apart hoofdstuk in mijn proefschrift, maar helaas hebben tijd en de ruimte mij ingehaald. Hier volgt een poging om mijn dank aan jullie uit te drukken.

Alles begint met de patiënten waarvoor dit onderzoek dient en zonder wie dit proefschrift nooit vorm had gekregen. De begeleiding van dit traject was in handen van dr. Paul van der Heijden, prof. dr. Roy Kessels en prof. dr. Jos Egger.
Paul, je kwam wat later, maar je kwam als geroepen. Heel concreet manuscripten vormgeven, maar ook vrijuit praten over wat het doen van patiëntgebonden onderzoek betekent voor ons werkveld was zeer waardevol en leuk. Dat je me behoedde voor 'Calimerodenken' (zaken niet te klein maken, red.) gaf me zekerheid. Je moet weten dat ik altijd positief en vol inspiratie terugkwam van onze afspraken in Den Bosch, Nijmegen of Venray. Jij laat onderzoek doen makkelijk lijken en creëert een leerzame en ontspannen samenwerking, waar ik je erg dankbaar voor ben.
Roy, jij was de grondlegger van mijn masterthese en stuurde me in de richting van het intelligentieonderzoek. Nooit vergeet ik meer dat je me in het Venloosch hebt toegezongen op mijn afstuderen (waarvoor dank, Jos), wat me vervolgens 100% vertrouwen gaf in een verdergaande samenwerking tijdens mijn PhD. Niet zelden hielp je me scherp te blijven als ik weer dreigde te vervallen in vage beschrijvingen en onnodige uitweidingen, waar ik je bijzonder voor wil bedanken. Ik heb bewondering voor je kennis, kunde, rust en werkgeheugencapaciteit.
Jos, mijn dank voor jou gaat veel verder dan je begeleiding bij het schrijven van dit proefschrift. Als een mentor heb je me begeleid vanaf het begin van mijn professionele carrière. Op de universiteit bij de sectie Klinische Psychologie, binnen het Topklinisch Centrum voor Neuropsychiatrie en met het onderzoek, overal heb je me laten zien hoe deze wereld werkt en me de ruimte gegeven om me laten ontdekken hoe ik me daarin beweeg. In het begin samen naast elkaar met de laptops op tafel in A.07.35 ('je moet me wel altijd in de cc zetten Loes!'), later wat meer op afstand door telefoontjes 's avonds op weg naar huis in de auto ('verder nog dingen?'). Je bent de meest drukbezette man die ik ken en ondanks dat mag en kan ik altijd alles aan je vragen. Gelukkig blijft het niet bij inhoudelijke zaken. Duitse gezegdes, Bijbelverhalen, geschiedenislessen, levenskwesties en vooral heel veel nieuwe woorden, jij als levende encyclopedie hebt mijn kennis aanzienlijk vergroot. Zou het dan toch, na mijn rijbewijs, nieuwe auto, nieuw huis, opleiding én proefschrift, eindelijk volwassen..?

William, jij was onmisbaar om mijn gedachtenkronkels over onderzoek in goede statistische banen te leiden. Ik ben je zeer erkentelijk voor je geduld, je kritiek, je tijd en je Nespresso.

Begeleiders kun je niet genoeg hebben en naast bovengenoemden kan wat vrouwelijke input ook zeer gewenst zijn. Lieve Ellen, hoewel jij vanaf de zijlijn betrokken bent geweest bij mijn begeleiding, ben je zo waardevol geweest tijdens het hele traject. Ik leerde je kennen als mijn stagebegeleidster en eigenlijk ben je dat nog steeds. Mijn sporadische 10-minuten gesprekjes met jou aan de telefoon of tijdens een ritje terug naar Nijmegen schepten direct orde in mijn chaos. Je hebt zo'n scherpe blik en hebt oog voor detail, maar je verliest nooit het doel uit ogen en hebt hart voor de patiëntenzorg. Velen hadden het vertrouwen dat dit me zou gaan lukken, maar als jij het zei kon ik het stiekem soms ook zelf geloven.
Cilia, toen het gevoel me bekroop dat ik verdwaalde in onderzoeksland, werd ik door jou geadopteerd. Ik ken niemand anders die zo'n rust uitstraalt in de hectische omgeving van de universiteit. Je bent enorm kritisch en stelt tegelijkertijd gerust, dat bewonder ik. Naast jou dank ik ook de DDM-club, waaronder Naline, Moniek, Leontien, Bea, Rose, Arie, Erwin en Ullrich voor hun input en constructieve ideeën en de schrijfweken. Ik wens jullie allen heel veel succes met al jullie projecten.

Ik dank de Raad van Bestuur van het Vincent van Gogh voor het faciliteren van wetenschappelijk onderzoek binnen de patiëntenzorg. Ik heb mijn onderzoek grotendeels verricht binnen het Topklinisch Centrum voor Neuropsychiatrie, waar expertise wordt ingezet en gedeeld ten behoeve van de zorg. Mensen zoals ik worden hier opgeleid tot kritische denkers, wat naast het uitpluizen van complexe diagnostische vraagstukken ook vaak leidt tot inspirerende discussies met collega's. Dat deed ik met Wouter, Linde, Laureen, Claudette, Annemieke en nog vele anderen. Liefste Wouter, zonder jou was dit proefschrift er natuurlijk niet. Sparringpartner for life, eerst in de trein, later in de auto, nu vooral via de mail of telefoon. De plannen voor ons eigen dynamische imperium om de wereld mee te veroveren blijven bestaan. Eerst nog even een opleiding of promotie hier en daar, maar eigenlijk zijn we er al klaar voor hè. Annemieke, dank voor de samenwerking en heel veel succes met jouw laatste loodjes! Linde, je bent een doorzetter en ik kijk trots vanaf de zijlijn naar jouw project! Laureen en Claudette, dank voor jullie hulp bij de dataverzameling en de nodige humor om het vol te houden! Renée, onze wegen liepen veelal parallel, daarom vind ik het leuk dat we nu samen de opleiding doen. Ruud en Henriëtte, jullie bedankt voor alle hulp de afgelopen jaren en dat jullie altijd de telefoon opnemen als ik bel!

De laatste loodjes van dit proefschrift schreef ik tijdens mijn opleiding binnen het Topklinisch Centrum voor Korsakov en alcoholgerelateerde cognitieve stoornissen, waar onderzoek en praktijk als vanzelfsprekend hand in hand gaan. Serge, ik heb nog even mogen afkijken bij jou, wat heb je het goed gedaan! Anke, wat heb ik in korte tijd ook van jou veel mogen leren, meer dan je denkt, dank daarvoor. Gwenny, jij was natuurlijk altijd mijn promotievoorbeeld, ook al wilde je dat niet. Ik vind het supertof dat we nu weer samenwerken. Ik bewonder jouw kracht en je relativeringsvermogen. En hoe je omgaat met je zelfgecreëerde chaos, want die is zo mogelijk nog groter dan die van mij.
Inspiratie vond ik ook bij mijn collega's van de sectie Klinische Psychologie aan de Radboud Universiteit, als ook mijn mede stagecoördinatoren Karin, Camille en Esther. Het geeft energie om zo'n gevarieerde en gedreven groep mensen te zien creëren en ontwikkelen. Amras, zo jong als we waren toen we begonnen zijn we inmiddels niet meer, maar we coördineren nog lekker even door toch? Liefst mét een chai latte. Heel veel succes met jouw onderzoek. Lieve Iris, wat ben ik blij dat jij er opeens was, op de uni én op mijn balkon (liefst mét een GT), ik hoop dat je nog even blijft. En dat je je ideeën op papier gaat zetten.

Excuses aan mijn vrienden is op zijn plaats. Lieve Katinka, Jo, Tom en Josje, ik hoop vanaf nu weer wat meer tijd te hebben om naar Hünxe te rijden (ik weet eindelijk de weg Kat!), te lunchen bij Bagels & Beans (*my god* Jo, zo lekker), eindelijk die volgende aflevering van *Shameless* te kijken (bij jou of bij mij Tom?) en te spontaan een avondje af te spreken (liefst mét borrel toch Josje?).

Een groot gedeelte van het schrijven van een proefschrift bestaat uit beseffen dat niet alles om het onderzoek draait. Bij mij werkte dat het beste door af te dalen naar het zuiden. Lieve Rens, lieve Renee, lieve Merel, Kessel is mijn toevluchtsoord waar mijn Nijmeegse leven niet bestaat en het is mijn thuis. Ik kijk uit naar alle vakanties en weekenden die ik nog met jullie ga doorbrengen. Nieuwe energie kreeg ik ook van mijn lieve Venlose meisjes en alle baby's die jullie hebben gemaakt. Ik heb ze allemaal even lief. Toiny, hoe anders lopen onze levens en hoe blij ben ik dat wij nooit voorbij zullen gaan. Jann, denk jij ook nog zo vaak terug aan onze reis? Zouden we dat ooit weer doen? Lieke, ik vind het supergaaf dat jij met je mooie en grappige hoofd de illustraties van dit boekje hebt bedacht en gemaakt, bedankt. Mamanda, ik huil nu waarschijnlijk net zo hard als jij, weet dat het je siert. Eefke, kleine meisjes doen blijkbaar hele toffe dingen, ik ben echt heel erg trots op jou.

Terug in Nijmegen heb ik altijd mijn paranimfen in de buurt. Onze overeenkomsten bleken veel breder dan alleen onze interesse voor de psychologie. Ik ben zo ongelofelijk blij en trots dat jullie nu naast me staan. Anne, jij kiest je eigen pad en daar heb ik heel

veel bewondering voor. Hoe jij denkt aan anderen, wil ik dat anderen denken aan jou. De raarste én grappigste dingen heb ik met jou meegemaakt. Maar dit is misschien toch wel het raarste wat we samen doen. Bij jou en bij Karst ben ik mezelf, is het altijd goed en voel ik me geborgen, of we nou bij elkaar wonen of niet. Laten we het altijd zo blijven doen, ok?
Suus, je bent in mijn leven gekomen en je mag nooit meer weggaan. Je kent al mijn kanten, soms beter dan ik ze zelf (wil) ken(nen). Ik ben blij met Stefan die het altijd goed vindt, hoe vaak we ook afspreken. En ik ben blij met Ruben, hij maakt je nog mooier dan je al was. Je bent niet te stoppen, zelfs niet als je heel soms zelf denkt van wel. Je bent van onschatbare waarde.

Lieve pap en mam, door jullie ben ik geworden wie ik nu ben. Het had ook heel anders, maar net zo mooi kunnen zijn, toch Sef? Pap, als ik met jou praat voel ik me gesterkt. Ik weet dat je altijd trots bent op mij, maar ik kan nu wel zeggen dat ik zelf ook trots ben op wat ik heb gedaan! Mam, wat kan ik zeggen, het boekje is eindelijk af, weer een zorg minder! Je bent er altijd en overal, of je er nu wel bent of niet, je stelt me gerust, waar we ook zijn en wat er ook gebeurt. Ik hou van jullie.

Curriculum Vitae

Curriculum vitae

Mrs. Loes van Aken was born in Venlo on August 9th, 1987. After completing her Athenaeum in 2005 (Blariacum College, Venlo), she started her study Psychology at Radboud University, Nijmegen, majoring in Neuropsychology and Rehabilitation Psychology. During her undergraduate years, Loes participated in the *Radboud University Honours programme*, which she completed in 2008. Her master thesis stimulated her interest for the relationship between fluid intelligence and executive functioning and later founded her PhD research that was performed at the Vincent van Gogh Centre of Excellence for Neuropsychiatry in Venray, where she also completed her clinical internship. In 2009, she obtained her master's degree cum laude and started working as a psychologist at the Vincent van Gogh Institute, where she continued combining her interests as a scientist and practitioner both by working as a clinician and starting with her dissertation research. As an external PhD student, embedded within the Donders Institute for Brain, Cognition, and Behaviour and the Behavioural Science Institute, both of Radboud University, Nijmegen, Loes also enrolled in the postgraduate residency program for health care psychologists. She combines her clinical and research activities with academic teaching as internship coordinator of Clinical Psychology for the Radboud University master program Mental Health Care Psychology for which she acquired her university teacher qualification in 2012 (BKO).

At present, she is active in the field of (neuro)psychodiagnostics and psychological treatment of patients with severe mental illness such as affective, anxiety and psychotic disorders, autism spectrum disorders, genetic syndromes with specific psycho-pathological phenotypes as well as Korsakoff syndrome and neurodegenerative disorders. She founded and chaired the Vincent van Gogh institutional committee on psychodiagnostic assessment, supervising psychologists, psychodiagnostic workers and interns for many years. Recently, she started as a clinical research teacher in the postgraduate residency program for clinical psychology specialists of Radboud Centre for Social Sciences.

Publicaties

Publications

International papers

Janssen, G., Van Aken, L., De Mey, H., Witteman, C., & Egger, J.I.M. (2013). Decline of Executive Function in a Clinical Population: Age, Psychopathology, and Test Performance on the Cambridge Neuropsychological Test Automated Battery (CANTAB). *Applied Neuropsychology: Adult, 21,* 210-219. doi:10.1080/09084282.2013.793191

Van Aken, L., Van der Heijden, P.T., Van der Veld, W.M., Hermans, L., Kessels, R.P.C., & Egger, J.I.M. (2015). Representation of the Cattell-Horn-Carroll theory of cognitive abilities in the factor structure of the Dutch-language version of the WAIS-IV. *Assessment,* 1-9. doi:10.1177/1073191115607973

Van Aken, L., Kessels, R.P.C., Wingbermühle, E., Van der Veld, W., & Egger, J.I.M. (2015). Fluid intelligence and executive functioning more alike than different? Latent variable analysis using KAIT and CANTAB. *Acta Neuropsychiatrica, 28,* 31-37. doi:10.1017/neu.2015.46

Van Aken, L., Kessels, R.P.C., Wingbermühle, E., Wiltink, M., Van der Heijden, P., & Egger, J.I.M. (2014). Exploring the Incorporation of Executive Functions in Intelligence Testing: Factor Analysis of the WAIS-III and Traditional Tasks of Executive Functioning. *International Journal of Applied Psychology, 4,* 73-80. doi:10.5923/j.ijap.20140402.05

Book chapter

Egger, J.I.M., Wingbermühle, E., & Van Aken, L. (2012). Cognitieve functiestoornissen. In Pepplinkhuizen, L., Verhoeven, W.M.A., & Sitsen, J.M.A. *Het psychiatrisch formularium; een praktische leidraad* (7e herziene versie, pp. 144-151). Houten: Bohn Stafleu van Loghum

Published abstracts

Egger, J., Janssen, G., Van Aken, L. Psychopathology and aging: executive function performance on the Cambridge Neuropsychological Test Automated Battery (CANTAB). *European Psychiatry, 29,* EPA-0533. doi:10.1016/S0924 9338(14)77929-5

Egger, J., Van Aken, L., Van der Heijden, P., Wingbermühle, E. & Kessels, R. (2014). WAIS-III taps fluid intelligence and executive functions in psychiatric patients and healthy participants. *European Psychiatry, 29,* EPA-0570. doi:10.1016/S0924 9338(14)77960-X

Egger, J.I.M., Verhoeven, W.M.A., Janssen, G.T.L., Van Aken, L. & Hoogeboom, A.J.M. (2012) Executive dysfunctions as part of the behavioural phenotype of Aarskog-Scott syndrome. *European Psychiatry, 27,* P-701. doi:10.1016/S0924-9338(12)74868-X

Egger, J.I.M., Van Aken, L., Kessels, R.P.C., Wingbermühle, E., Van der Veld, W. & Verhoeven, W.M.A. (2011). Fluid intelligence and executive functioning: Partial overlap in patients with psychiatric disorders. *European Psychiatry, 26,* S1:P03-46. doi:10.1016/S09249338(11)72920-0

Van Aken, L., Egger, J.I.M., Wingbermühle, E., Van der Veld, W., Kessels, R. (2010). Fluid intelligence and executive functioning: A latent variable analysis of the Kaufman Adult Intelligence Test (KAIT). *Abstracts of the Second Meeting of the Federation of European Societies of Neuropsychology (ESN),* 155-155.

Donders Series

Donders Graduate School for Cognitive Neuroscience

For a successful research Institute, it is vital to train the next generation of young scientists. To achieve this goal, the Donders Institute for Brain, Cognition and Behaviour established the Donders Graduate School for Cognitive Neuroscience (DGCN), which was officially recognised as a national graduate school in 2009. The Graduate School covers training at both Master's and PhD level and provides an excellent educational context fully aligned with the research programme of the Donders Institute.

The school successfully attracts highly talented national and international students in biology, physics, psycholinguistics, psychology, behavioral science, medicine and related disciplines. Selective admission and assessment centers guarantee the enrolment of the best and most motivated students.

The DGCN tracks the career of PhD graduates carefully. More than 50% of PhD alumni show a continuation in academia with postdoc positions at top institutes worldwide, e.g. Stanford University, University of Oxford, University of Cambridge, UCL London, MPI Leipzig, Hanyang University in South Korea, NTNU Norway, University of Illinois, North Western University, Northeastern University in Boston, ETH Zürich, University of Vienna etc.. Positions outside academia spread among the following sectors: specialists in a medical environment, mainly in genetics, geriatrics, psychiatry and neurology. Specialists in a psychological environment, e.g. as specialist in neuro-psychology, psychological diagnostics or therapy. Positions in higher education as coordinators or lecturers. A smaller percentage enters business as research consultants, analysts or head of research and development. Fewer graduates stay in a research environment as lab coordinators, technical support or policy advisors. Upcoming possibilities are positions in the IT sector and management position in pharmaceutical industry. In general, the PhDs graduates almost invariably continue with high-quality positions that play an important role in our knowledge economy.

For more information on the DGCN as well as past and upcoming defenses please visit:
http://www.ru.nl/donders/graduate-school/phd/